INDUSTRIAL PSYCHOLOGY

INDUSTRIAL PSYCHOLOGY

by

JOSEPH TIFFIN, Ph.D.

PROFESSOR OF INDUSTRIAL PSYCHOLOGY
PURDUE UNIVERSITY

Second Edition

New York
PRENTICE-HALL, INC.
1947

First Edition

First Printing............June, 1942
Second Printing.......October, 1942
Third Printing.......February, 1943
Fourth Printing........August, 1943
Fifth Printing..........March, 1944
Sixth Printing........February, 1946
Seventh Printing.......August, 1946
Eighth Printing........October, 1946

Second Edition

First Printing............July, 1947

PRINTED IN THE UNITED STATES OF AMERICA

TO

F. B. K.

H. S. K.

AND

R. A. S.

Preface to the Second Edition

The five years since the publication of the first edition of *Industrial Psychology* have been eventful ones for the world and for psychology in industry. A decade ago, only a dozen or fifteen persons were usually in attendance at sessions on industrial psychology at professional meetings. Hundreds now attend the sessions on this subject. A decade ago, only a scattered few psychologists were employed by industry. Scores of psychologists are now well established in industrial jobs. Before World War II, only a few courses in industrial psychology were given by a limited number of colleges and universities. Now, almost every department of psychology gives one or more courses in this subject, to say nothing of the many extension, evening, and correspondence courses devoted to psychology in industry.

This growth of interest in the field of industrial psychology has also resulted in a marked increase in publications in this field during the past few years. The major purpose of the present revision of *Industrial Psychology* has been to include a coverage of this extensive new material. Every chapter has been thoroughly revised and much new material has been included. In addition, two entirely new chapters have been added, one of which deals with the interview and related employment procedures, the other with wages and job evaluation. The recent recognition of the contribution that can be made by psychology to the subject of job evaluation would seem to justify a coverage of at least the basic principles of job evaluation techniques in a textbook on industrial psychology. The chapter on visual skills has been revised in collaboration with Dr. S. E. Wirt.

To an even greater extent than when the first edition was written, the author is indebted to many persons in the industrial world for continued interest in and support of this work. Indeed, the number of persons and companies to whom appreciation is due is so great that even a sincere attempt to list them would almost certainly result in some unintended omissions. Specific acknowledgments to those from whom work has been reported are given throughout the text. Other than this, no complete list of acknowledgments will be attempted, but a specific word of appreciation is clearly due to the Bausch and Lomb Optical Company for its collaboration and material assistance in research on personnel testing procedures. This company has made possible the installation and operation at Purdue University of a statistical research laboratory—the Occupational Research Center—, under the supervision of Dr. N. C. Kephart and Dr. S. E. Wirt. Without this coöperative assistance, many of the investigations reported in this revision could not have been made.

JOSEPH TIFFIN

Preface to the First Edition

THIS book deals with applications of psychology that have been made in industry. These applications are not limited to employee selection and placement. Industrial psychology has also been applied to the improvement of merit rating, reduction of accidents, solution of visual problems, increasing the accuracy of inspection, improvements in training methods, and the measurement and improvement of employee morale.

The growth of interest in psychological methods during the past decade within such organizations as the American Management Association shows that psychology as a *technology* has been accepted as a tool of industrial management. This book covers the procedures and techniques that have been responsible for that acceptance. These techniques can be improved, of course. They will be improved as further industrial application of psychology points the way toward desirable modifications. But just as they are *now*, the industries that have given them a trial have not been disappointed in them.

A considerable amount of the content of this book is based on research that has not previously been published. This work has been in the nature of coöperative projects carried on by the Division of Education and Applied Psychology at Purdue University and a number of industries. Practically all of the material on industrial vision and accidents, and a considerable amount of the material on individual differences, merit rating, and employee placement tests is presented here for the first time.

The treatment of test application emphasizes the importance of the selection ratio—an emphasis first crystallized

by H. C. Taylor and J. T. Russell of the Western Electric Company. Test programs conducted in terms of the principle have shown that a highly effective use of tests can be made even when *no* applicants are rejected for employment. The value of a testing program under such conditions has not always been recognized. To facilitate an understanding of the selection ratio among students and to make available to personnel men the use of this important procedure, the Taylor-Russell Tables have been reproduced in Appendix B, with the permission of Dr. Taylor.

The writer is indebted to a number of persons and organizations for assistance in this work. F. B. Knight, Director of the Division of Education and Applied Psychology at Purdue University, has constantly encouraged and assisted the research in industrial psychology. R. J. Greenly, Professor of Trades and Industrial Education at Purdue University, has helped in numerous ways to apply psychology to industrial problems. The Bausch and Lomb Optical Company has assisted and supported research dealing with the construction and validation of a battery of vision tests adapted to the needs of industry. Hedwig S. Kuhn, M. D., furnished more than 9000 visual records from an industrial survey. Most of the vision research is based upon an analysis of these records. The joint committee on industrial vision of the American Medical Association and the American Academy of Ophthalmology and Otolaryngology has given valuable counsel in connection with the vision research. The industrial vision committee of the American Optometric Association has also been enthusiastic in its interest and support.

Among the many men in industry who have given helpful and sound counsel on the industrial applications of psychological methods the writer is particularly indebted to Yandell Cline and Paul Ortlieb, of Noblitt-Sparks Industries, Inc.; to I. P. Egan, of the Real Silk Hosiery Mills; to William Pope, of the Bear Brand Hosiery Mills; to E. D. Stoetzel, W. E. Strong, and Rei Ganzer, of the Marathon Paper

Company; and to F. J. Martin, of the Belden Manufacturing Company.

Most of the research on industrial vision has been carried on in collaboration with S. E. Wirt and Ray Reed, of the Bausch and Lomb Optical Company. Chapter 6, which deals with this subject, has been written in collaboration with Dr. Wirt. Collaboration with H. B. Rogers, Associate Professor of Industrial Engineering at Purdue University, in research on the selection and training of industrial inspectors, has resulted in a relationship between psychology and industrial engineering that has broadened the horizon of applied psychology. The interest in industrial psychology constantly expressed by Dr. Lillian Gilbreth, Professor of Management at Purdue University, has been a source of inspiration. Mr. G. A. Satter read and suggested certain changes in the appendix dealing with elementary statistical procedures. Mr. Max Wastl, laboratory technician, constructed most of the apparatus used in instrumental research.

To all of the above, to the numerous workers whose published research has been drawn upon in writing this book, and to the other persons in industry who have helped with various phases of the research in industrial psychology, the writer is happy to acknowledge his sincere appreciation.

JOSEPH TIFFIN

Contents

APPENDICES

1

The Significance of Individual Differences in Industry

THE management of an expanding industry may obtain ten or a hundred or even a thousand new machines of a certain type with reasonable assurance that the new pieces of machinery will be identical in construction, equal in efficiency, and capable of uniform output. But to obtain an equal number of persons who are able to operate these machines with a satisfactory and uniform degree of competence is quite another matter. Every foreman knows that identical machines seldom deliver identical production when they are controlled by different operators. People are not alike by nature, training, education, or inclination. A job may be done well by one, fairly well by another, and very poorly by a third. The job may be a source of great personal satisfaction to one, a monotonous and boring task to another, and entirely beyond the capacity of a third. Modern industry is becoming increasingly aware of the importance of placing on every job an individual who is not only able to *do* the job well but who, in addition, is temperamentally adapted to the job in question. The success of personnel placement depends upon placing every individual on a job that matches the capacity of the individual and upon giving him adequate and specific training to do the job. If the job is too difficult or the training has been inadequate, the result is confusion, low production, and possible injury either to the operator or the machine. If the job is too easy, the result is boredom, mind wandering,

and daydreaming, with the dissatisfaction that so often accompanies these activities.

An individual is best adapted and is usually most satisfied —and this applies whether he is at work, at home, or on his vacation—when he has found an outlet for whatever energy, drive, and ability he may possess. If his job calls for abilities that he does not have and cannot develop, he continually experiences the despair of failure. If, on the other hand, his job calls for only a fractional part of his ability, he is likely to develop other means of self-expression, which, at their best, may be daydreaming or an unduly critical attitude and, at their worst, may become a definite and serious mental illness or disease. Energy not demanded by the job is usually released into some other channel. Too often the release of this "extra energy" is in a direction that not only fails to benefit either the employee or his employer but, on the contrary, is actually detrimental to the interests of one or both of these parties.

Many investigations, both experimental and statistical, furnish the basis for the above statements. In 1923, Bills[1] reported a significant correlation between intelligence and difficulty of work being done by clerical workers. She noted that after the workers had been two and a half years on the job, the correlation approximately doubled in amount. This she interpreted to mean that with seniority comes a definite shifting of employees toward levels of job difficulty that match their ability. In another investigation, Pond and Bills[2] found that labor turnover can be markedly reduced by a careful placement of employees in jobs of a difficulty commensurate with the ability of those employees. Low-ability employees—as revealed by mental tests at the time of employ-

[1] M. A. Bills, "Relation of Mental Alertness Test Scores to Positions and Permanency in Company," *Journal of Applied Psychology*, VII (1923), pp. 154–156.

[2] Millicent Pond and Marion A. Bills, "Intelligence and Clerical Jobs. Two Studies of Relation of Test Score to Job Held," *Personnel Journal*, XII (1933), pp. 41–56.

ment—showed only half the turnover on certain simple jobs as did high-ability employees assigned to the same jobs. On jobs more difficult in nature, exactly the opposite situation prevailed: the low-ability employees showed the largest turnover and the high-ability employees the least. Standardized mental tests[3] reveal that on jobs of a repetitive, routine nature it is not uncommon to find negative correlations between productivity and mental ability. Such studies indicate the *practical* importance of placing employees on jobs that match their ability.

Industry has long recognized the existence and importance of individual differences in training and skill. Personnel managers, in hiring tradesmen, make every effort to determine in advance the degree of skill that the applicant possesses. But often employers do not so clearly recognize the fact that differences in capacity for machine operation and other jobs that are ordinarily considered as unskilled or semiskilled are just as great and just as important as differences in skill among tradesmen. The significance to industry of individual differences among employees is far more important than that of differences in skill already developed or of differences that can be detected in an interview. The concept of individual differences is concerned with basic differences in capacity which are of importance in every phase of industrial personnel placement.

Recent social legislation and the inclusion of some form of seniority clause in practically every union contract make the consideration of individual differences at the time of hiring more important than the placing of men and women on jobs that are neither above nor below their capacity to succeed. There was a time, not so long ago, when any employee—new or old—could be dismissed whenever his supervisor felt that

[3] Joseph Tiffin and R. J. Greenly, "Employee Selection Tests for Electrical Fixture Assemblers and Radio Assemblers," *Journal of Applied Psychology*, XXIII (1939), pp. 240–263. See also Joseph Tiffin and C. H. Lawshe, Jr., "The Adaptability Test: A Fifteen Minute Alertness Test for Use in Personnel Allocation," *Journal of Applied Psychology*, XXVII (1943), pp. 152–163.

the employee's services were no longer advantageous to the company. Often employees were tried out on several jobs over a period of months and were then dismissed as unsuited for any of the jobs available. Under modern legislative enactments and union-management agreements, it is practically impossible for industry to follow this procedure today. Unemployment insurance premiums are paid at least partly by the employer, and the schedule of premiums is so adjusted in many states that the amount which an employer pays is proportional to the labor turnover of the company. Union contracts often make it difficult to discharge an employee after he has been employed for a specified period of time.

At least two other factors operate to prevent the arbitrary dismissal of employees. One of these is based upon financial considerations, for the cost of training an employee is lost when that employee is dismissed. Training costs are no small item in the operation of a modern plant—a matter that will be discussed in some detail in Chapter 8. By using every precaution to make sure that training is given only to those employees who are likely to stay on the job, at least until the investment in their training has been realized, and by making every effort to avoid the dismissal of these employees for *any avoidable reason*, sizable savings in production costs can almost invariably be made.

A second less mercenary, but very real, reason for a growing reluctance on the part of management to dismiss employees arbitrarily is the increasing sense of social obligation felt by a very large number of employers. At one time management regarded labor simply as a cost of production, and *only* as a cost of production. Today management recognizes that employees not only *make* the product, but also *buy* it; and that employees who leave dissatisfied may, and frequently do, spread discontent among other employees or future employees that works to the disadvantage of both producer and consumer. Satisfactory employer-employee relations are the cornerstone upon which a modern industrial organization

must be built. For such reasons it has become more and more important for industry to evaluate the suitability of prospective employees *before hiring.* Differences in the suitability of applicants is a branch of the psychology of individual differences that, fortunately, has been quite thoroughly studied. To show the direct and practical application of these studies in the field of personnel administration will be a major purpose of this book.

Several questions concerning individual differences have probably already arisen in the mind of the reader. How great are individual differences? Are they large enough to be of practical importance? Do they indicate more or less permanent characteristics? How are they affected by experience and/or training on the job? If employees were given equal experience, would not their individual differences largely disappear? Questions of this type can be answered most satisfactorily by consideration and interpretation of statistical data and experiment.

The Magnitude of Individual Differences

Individual differences in productivity

Figure 1 is a frequency distribution showing differences in productivity among thirty-six electrical-fixture assemblers who were engaged in identical jobs. An incentive method of wage payment was in use and the best operator was earning slightly more than twice the average hourly wage of the poorest operator. The overhead to the company was identical for all the employees in this department, including the best and the poorest. Both used identical work layouts and assembling equipment. Both used identical amounts of space, heat, light, and other overhead expenses. But the best employee, in delivering more than twice the production of the poorest employee, was therefore using only half as much overhead and capital expense *per unit of production* as the poorest employee. This fact makes it clear that as new

employees are added to the department it is profitable not alone to the employees or to the management but *both* to the employees and the management, to obtain persons who are capable of attaining the higher levels of production.

Figure 2 shows another distribution of the productivity of a group of operators on identical jobs (burning, twisting, and soldering) utilizing identical investments in machinery and overhead. The distribution shows that the production

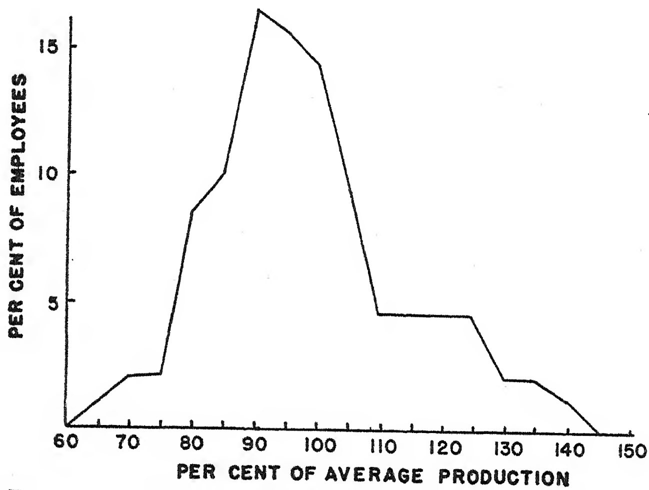


Fig. 1—Distribution of quantity of production among 36 electrical fixture assemblers.

obtained from the different identical sets of equipment varied in the ratio 65:130, *depending on the operator of the machine.*

Still more striking differences in productivity among a group of operators are shown in Figure 3. This distribution shows the production of 199 hosiery "loopers." The looping operation involves the gathering together of the loops of thread (each over a separate needle) at the bottom of a stocking after the garment is knit in order to close the opening left in the toe. Looping is a job calling for very careful and constant visual attention as well as a high degree of skill in placing the loops on the separate needles. Production on

this job is ordinarily computed in terms of dozens of pairs looped per hour. Payment is on an incentive basis.

Figure 3 shows a range of from .25 dozen pairs per hour (at the extreme left) to 7.00 dozen pairs per hour (at the

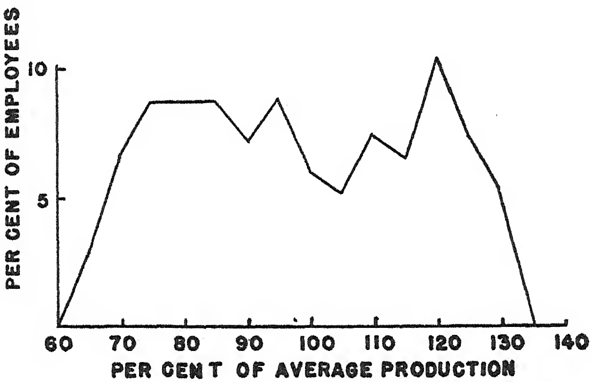


Fig. 2—Distribution of quantity of production among 33 employees engaged in an operation involving burning, twisting, and soldering the ends of insulated wire.

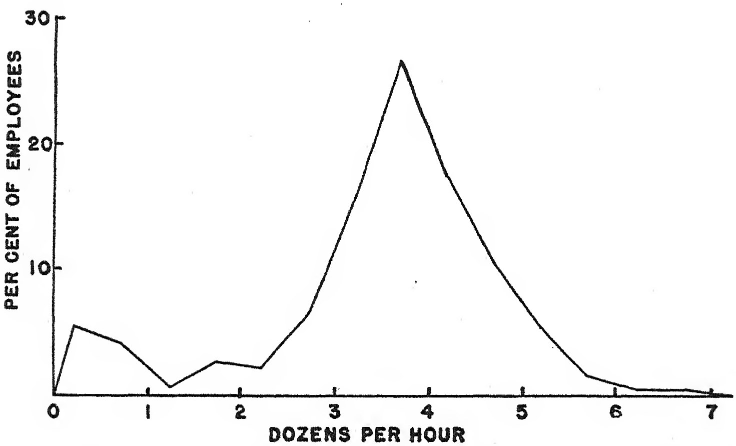


Fig. 3—Distribution of quantity of production in looping hosiery for 199 employees varying in experience from one month to five years.

extreme right). As might be expected, this great variation in production is due partly to differences in experience on the job. However, the fact that such differences do not account

for all of the variability shown in Figure 3 is indicated by Figure 4, which shows a similar frequency distribution of production of 99 employees who had had one year or more of experience on the looping job. A careful analysis of the learning curve for looping shows that maximum production is reached after a year of experience. It therefore seems reasonable to assume that the differences in production shown

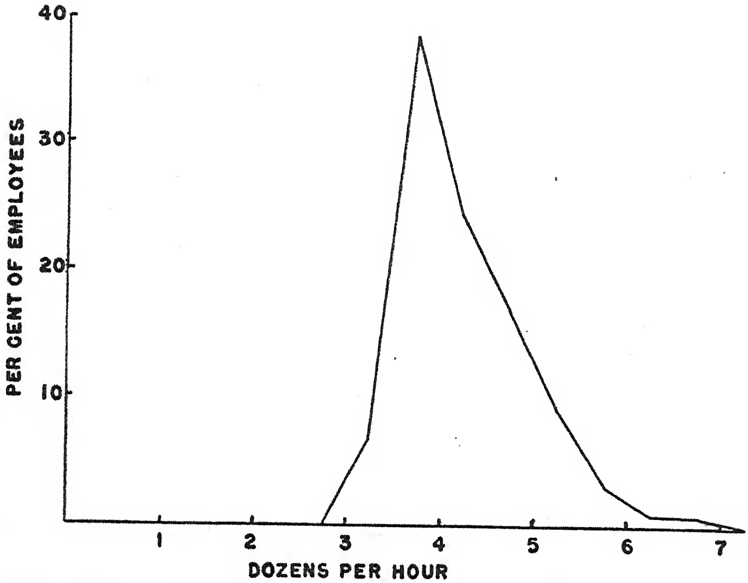


Fig. 4—Distribution of quantity of production in looping hosiery for 99 employees with one year or more of experience.

in Figure 4 cannot be explained in terms of the experience factor. Yet the variation in production is still from three dozen pairs to more than seven dozen pairs per hour.

Differences in employee productivity of the type just discussed are usually consistent differences. The high-producing employee at any one time tends to remain in the high-production level, and the less efficient operator tends to remain at approximately the same low level from week to week. This fact is illustrated by the scattergram shown in

Table 1. This table shows the relation between productivity for two successive weeks among 203 hosiery loopers. No marked shifts in productivity from one week to the other occurred. It is therefore clear that, given the production of any operator for the first week, a fairly accurate prediction of the production of that operator for the following week can be made. The coefficient of correlation[4] between the production of the first and second week computed from the scattergram in Table 1 was .96.

These computations indicate that an employee's production level is not something that fluctuates willy-nilly or that he changes daily, as he does his necktie or his shirt, but that his production level is rather a relatively fixed and permanent characteristic.

TABLE 1

SCATTERGRAM SHOWING CONSISTENCY OF PRODUCTION OF 203 HOSIERY LOOPERS

The correlation between production for the two weeks shown was .96

Average dozens per hour (second week)

Average dozens per hour (first week)	0–.4	.5–.9	1.0–1.4	1.5–1.9	2.0–2.4	2.5–2.9	3.0–3.4	3.5–3.9	4.0–4.4	4.5–4.9	5.0–5.4	5.5–5.9	6.0–6.4	6.5–6.9
6.5–6.9														1
6.0–6.4													1	
5.5–5.9											1	4		
5.0–5.4										1	11	1		
4.5–4.9								3	3	16	1			
4.0–4.4								11	16	3	1			
3.5–3.9							2	43	8	2				
3.0–3.4						1	25	3						
2.5–2.9						12	2							
2.0–2.4				1	3	1		1						
1.5–1.9				3	1									
1.0–1.4			2	1										
.5–.9		6	1											
0–.4	10	1												

[4] The meaning of a coefficient of correlation is explained in Appendix A, p. 511.

Individual differences in productivity, then, are large enough to warrant careful attention. This truth may be more easily understood if the fact, often lost sight of in the rush of industrial production, is kept in view that differences in productivity of employees are, basically and from a cost angle, *differences in overhead expenses and capital investment.*

Individual differences in job qualifications

Before employees have had an opportunity to reveal their differences in productivity on the job (that is, before they are hired) they will often, if given an opportunity, exhibit very great differences in qualifications for the job. A large manufacturing industry recently examined 112 applicants for the

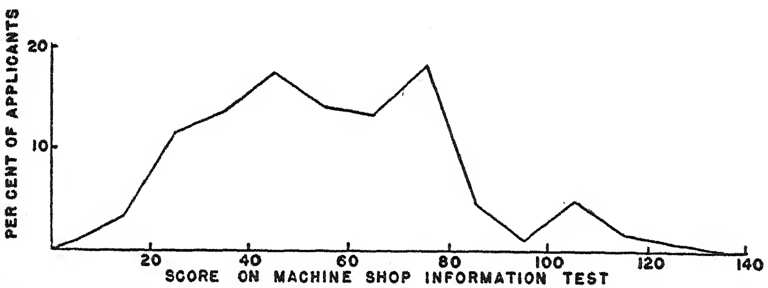


FIG. 5—Distribution of scores of 112 applicants on 130-item multiple-choice test covering technical information in machine shop practice.

job of machine-shop apprentice. Several tests were given to the applicants, among which was the Technical Information in Machine Shop Test of the Purdue Vocational Series.[5] Figure 5 shows that the scores on this test varied from 5 items correct to 125 items correct. All of the applicants either had taken high-school vocational courses, which presumably had prepared them in the technical aspects of machine-shop

[5] H. F. Owen, C. C. Stevason, H. G. McComb, and C. D. Hume, *Technical Information Test for Machinists and Machine Operators*, (Chicago: Science Research Associates, 228 South Wabash Avenue, 1942). This test is described briefly on p. 155. The original form of this test, on which Fig. 5 is based, contained 130 items. The revised form distributed by Science Research Associates contains 133 items.

practice, or had had industrial experience as machinists' helpers. In spite of these opportunities to learn the technical details of their craft, the test revealed enormous differences among the applicants in actual qualifications for the apprenticeship openings.

A similar situation among applicants for electrical apprenticeships is revealed in Figure 6. In this instance, the test

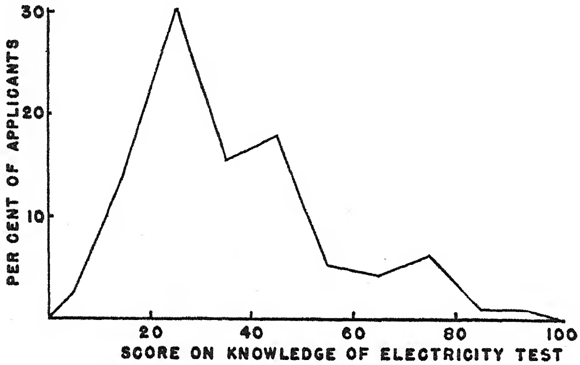


FIG. 6—Distribution of scores of 104 applicants on 100-item multiple-choice test covering knowledge of practical electricity.

covering Technical Information Related to Electricity of the Purdue Vocational Series[6] was used. Here again marked individual differences in qualification for the apprenticeships were revealed.

After the wide range in scores on the qualification tests described above was revealed, it was decided to administer the Otis Self-Administering Test of Mental Ability,[7] a standard intelligence test described briefly on page 83, as a means of obtaining additional information about the appli-

---

[6] C. W. Caldwell, H. R. Goppert, H. G. McComb, and W. B. Hill, Technical Information Series, Test for Electricians (Chicago: Science Research Associates, 228 South Wabash Avenue, 1942). This test is described briefly on p. 157. The original form of this test, on which Fig. 6 is based, contained 100 items. The revised form distributed by Science Research Associates contains 65 items.

[7] Arthur S. Otis, Otis Self-Administering Tests of Mental Ability (World Book Company, 1922).

cants in these two groups before a selection was finally made. The distributions of the intelligence-test scores for the two groups are shown in Figures 7 and 8. Many psychologists will find it hard to believe that high-school graduates of vocational courses tested as low as did those individuals at

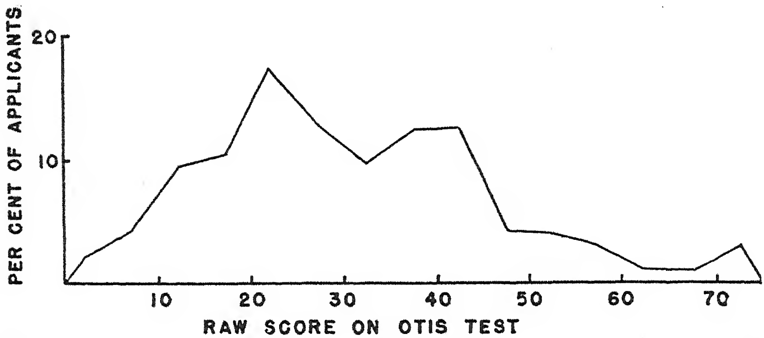


FIG. 7—Distribution of scores on Otis Self-Administering Test of Mental Ability (Higher Exam, Form A) of 112 applicants for machine shop apprenticeships.

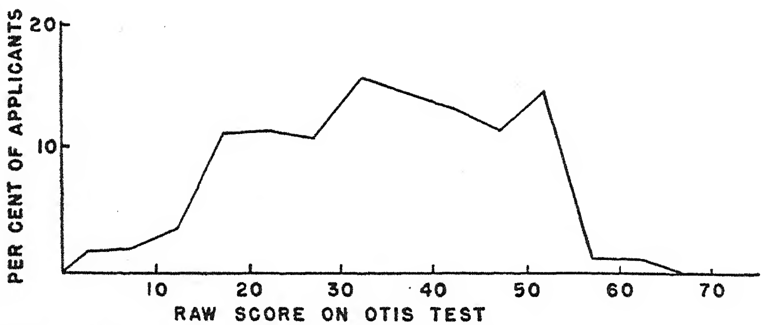


FIG. 8—Distribution of scores on Otis Self-Administering Test of Mental Ability (Higher Exam, Form A) of 117 applicants for electrical apprenticeships.

the left end of the scale in Figures 7 and 8. However, the distributions represent the scores exactly as they were obtained. It is very doubtful whether a boy with a raw score of 15 or less on the Otis test would be able to profit from an apprenticeship in machine shop or electricity to a sufficient extent to make the venture worthwhile either to the boy or to the company.

Job qualification tests are not limited to tests of technical information or intelligence. Many, if not most, jobs of a routine nature demand capacity for dexterous and co-ordinated activity more specifically than they demand technical information and mental ability. Just as one hundred persons selected at random would hardly be likely to possess equal capacities to become 100-yard-dash men, so one hundred persons selected by interview or otherwise would not be likely to possess equal capacity for a given production job *unless capacity for the production job is measured before they are hired.*

This fact is illustrated in Figure 9, which shows frequency distributions of four different groups[8] on the O'Connor test of finger dexterity.[9] This test consists of a 100-hole peg-board which is filled as rapidly as possible (three pins to a hole) from the pins located in a shallow tray. The greater one's finger dexterity, the more rapidly he will be able to fill the board and the smaller will be his score, which is simply the time in minutes required to fill the board. The original method of scoring this test, as described by Hines and O'Connor,[10] was somewhat more complicated than the simple determination of time in minutes required to fill the board. However, since Tiffin and Greenly[11] found a correlation of .99 between the simple time score and the scores obtained by the original formula, it was not necessary to compute the scores by the original formula. In Figure 9, the left side of the scale represents the small-time values or good scores, and the right side the long-time values or poor scores. The upper curve, or *A*, in Figure 9 shows the distribution of finger dexterity—as measured by this test—among a very large number of randomly selected persons. The remaining distributions in this figure (*B*, *C*, and *D*) show how finger

---

[8] Tiffin and Greenly, *op. cit.*

[9] M. Hines and J. O'Connor, "A Measure of Finger Dexterity," *Personnel Journal*, IV (1926), pp. 379–382.

[10] *Ibid.*

[11] Tiffin and Greenly, *op. cit.*

dexterity is distributed among three groups of employees engaged in work for which finger dexterity is presumably an important qualification. A marked similarity among the four distributions is revealed. Operators on the three

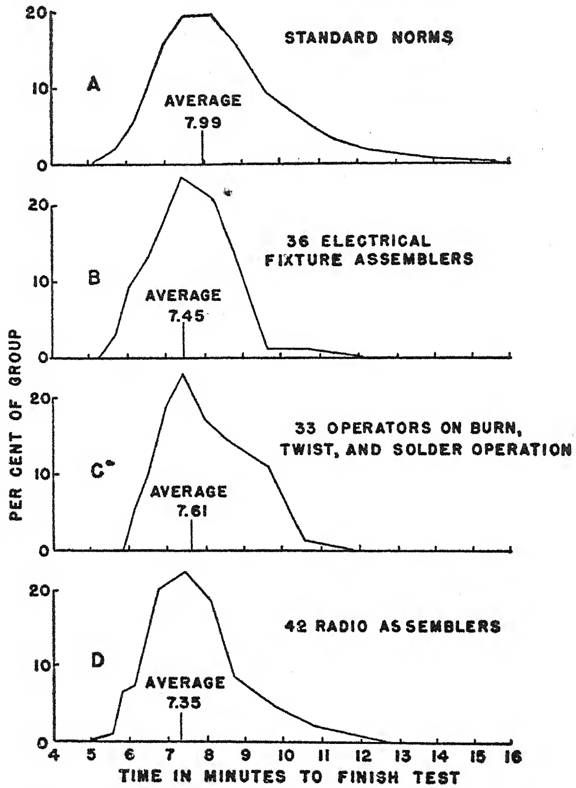


FIG. 9—Frequency distributions of scores on O'Connor Finger Dexterity Test of random group of subjects (top) and three groups of industrial employees engaged in rapid finger work.

industrial jobs are little better in finger dexterity than are randomly selected persons, which means that persons equally qualified for this job would have been obtained if the employees had been selected by lot from a telephone directory.

The question will naturally be raised: What of it? How

do we know that the finger-dexterity pegboard is a job qualification test for these jobs? Are the persons who test high in finger dexterity better employees than those who test low?

Detailed answers to these questions must be postponed at this point (they will be given in Chapter 5); but we may definitely say here that the test in question does measure something that is necessary for efficient production on the jobs under consideration, and we may add that the employees

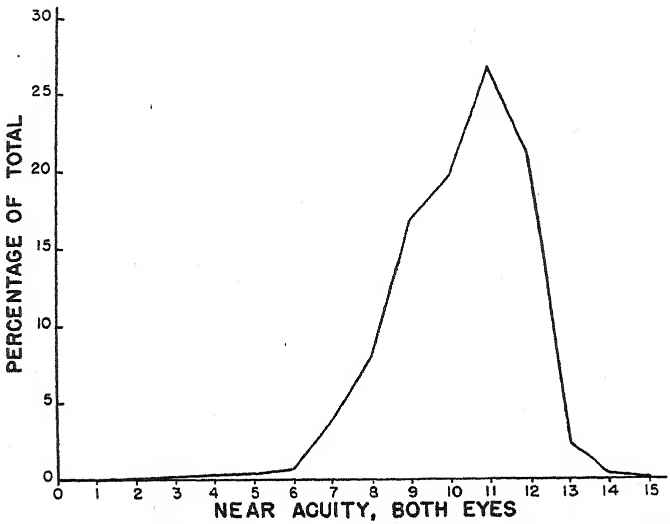


FIG. 10—Distribution of scores of 3,397 factory workers on a visual acuity test at 13 inches.

testing high on the finger-dexterity test are, in general, those employees who are also most productive at their work.

Differences in vision furnish a further example of individual differences in job qualification. Figure 10 shows a distribution of degree of visual acuity at a distance of 13 inches for 3,397 employees in a manufacturing company. While this curve reveals that a large proportion of these employees have "normal" acuity (10) or better, it also shows a definite percentage with vision below this level. It further reveals a scattering of employees with acuity so low that they

are, on certain types of jobs, severely handicapped at their work. This statement is based upon investigations showing the relationship between vision and efficiency (see page 199). Visual acuity is not the only aspect of vision in which individual differences among employees exist. Numerous tests dealing with other aspects of visual performance and efficiency reveal equally marked variations. A detailed treatment of vision as it relates to employee efficiency on different kinds of work will be covered in Chapter 7.

Merit ratings furnish another indication of individual differences in employee qualifications for certain jobs. A merit rating is a periodic judgment of an employee by his supervisor. It consists, usually, of a blank which permits separate ratings of the employee in terms of such factors as safety, productivity, industriousness, initiative, and so on. A survey in 1939 indicated that at that time approximately one-third of American industries were using some type of formal merit-rating system.[12] During the past few years many more plants have installed merit rating. Merit-rating systems, which will be discussed in detail in Chapter 10, usually result in a single over-all rating for each employee. This over-all rating may be taken, with certain qualifications, to indicate the employee's suitability on the job.

A distribution of the over-all ratings of 710 men from one department of a steel mill is shown in Figure 11. Many cautions should be observed in interpreting the significance of these merit ratings, but it is safe to conclude that the spread of over-all ratings from 240 to 450 points suggests definite differences in the quantity and quality of service rendered by the different employees.

Many additional instances of individual differences which affect industrial production could be cited, and all of them would point toward the conclusion that on many jobs a good

---

[12] R. B. Starr and R. J. Greenly, "Merit Rating Survey Findings," *Personnel Journal*, XVII (1939), pp. 378–384.

employee is at least twice as valuable to the company for which he works as a poor employee. We may answer the question about the magnitude of individual differences, then,

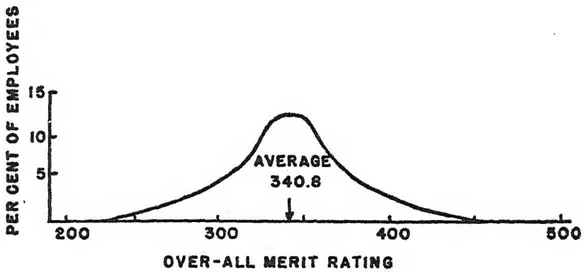


FIG. 11—Distribution of over-all merit ratings of 710 men from one department of a steel mill.

by saying that the differences are not only real but are large enough to be of vital practical importance to industry.

The Effect of Training Upon Individual Differences

Many psychologists have attempted to reach a general conclusion from experimental investigations of the effect of training upon individual differences. Three possible conclusions might be considered: training may increase the differences, it may decrease them, or it may leave them unaffected.

Examination of experimental studies in this field reveals what seems to be a good deal of contradictory evidence. In certain kinds of situations it seems clear that training tends to increase whatever differences in ability may exist at the beginning of the training. A group of persons placed upon an entirely new task of a complicated nature are not likely to differ much among themselves in their ability to handle that task at the beginning. Their skill is likely to be quite uniform and, it might be added, uniformly low, until appropriate training has given those with capacity to do the task an opportunity to forge ahead. The operation of any complicated industrial production machine will furnish an example of this principle.

In Figure 12 the heavy black line shows the average production of hosiery loopers in dozens of pairs per hour plotted against the amount of experience on the job. This curve is really the learning curve for this operation, even though it was obtained by taking the average production of loopers of varying amounts of experience rather than by following the improvement in skill of a given group of loopers as experience increased. The dotted lines in Figure 12 indicate a standard deviation[13] above and below average produc-

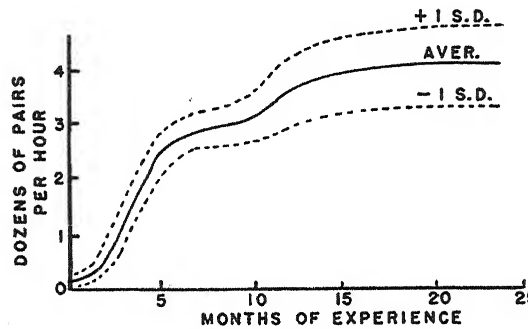


FIG. 12—Relation between experience and production in looping hosiery, and the effect of experience on individual differences in productivity. Results based on 199 loopers of varying experience.

tion for the varying amounts of experience. It is clear that the variability is greater among the loopers with considerable experience on the job than among those who have been recently hired or who are in the early stages of learning the operation. Here, then, is a specific case in which training *increases* the magnitude of individual differences with performance upon a given job. The operators are much more nearly like one another in the early stages of their training than in the intermediate or later stages.

One should not infer from the foregoing illustration, how-

[13] The standard deviation, a measure of variability or spread of scores, is explained in Appendix A, p. 498.

ever, that the effect of training is always to increase the magnitude of individual differences. For example, Hartmann,[14] using four tests of manual ability, found that three months of training resulted in definitely more improvement among those who initially tested low than among those who initially tested high. The ability of everyone was improved as a result of the training, but, in this experiment, the training also tended to decrease the difference between the good and the poor scores.

It would be possible to cite numerous other experiments that support the apparently contradictory conclusions of one or the other of the investigations cited above. A study of the experiments in question, however, suggests a possible explanation for the seeming discrepancy in results. Wherever the effect of training has been to increase the magnitude of individual differences, a study of the task involved usually reveals it to be fairly complicated. By complicated we mean that it is one in which the average individual does not reach his maximum level of performance without a fairly long learning period. Wherever training results in a decrease in individual differences, the task is usually found to be a rather simple one, that is, one upon which the average individual shows little improvement after a relatively short practice period. In the light of these facts we may venture the tentative generalization that training tends to increase individual differences in proportion to the complexity of the task in question. Although a few minor exceptions to this principle may come to mind, it seems to be at least provisionally acceptable as a working hypothesis.

Whatever may be the effect of training on the magnitude of individual differences, it seems clear that training seldom changes the relative standing of individuals in their ability to perform any given task. This conclusion is in accord with the

---

[14] G. W. Hartmann, "Initial Performance as a Basis for Predicting Ultimate Achievement," *School and Society*, XXIX (1929), pp. 495–496.

findings of Bishop,[15] Peterson,[16] Hartmann,[17] Viteles,[18] and others. No industry, therefore, should expect its training program to bring all employees up to the same high level of efficient performance unless the employees have initially been selected in terms of their capacity to reach that high level.

The Bases of Individual Differences

Psychologists have long been interested in determining the ultimate cause of individual differences among people. Usually they have divided the major causes into the two general categories of heredity and environment. Those upholding one point of view have often minimized or even completely ignored the possible influence of the other factor. It seems probable that both factors are usually operative but that their relative importance differs markedly in the determination of different personal characteristics. Heredity would seem to be of most importance in determining such traits as height, weight, and strength, although it is clear that environment also has some effect upon these factors. The environmental factors of training or education are of basic importance in determining trade skills, or any other type of skill that is reached only after experience on a particular job. But to conclude that differences in skill are dependent upon differences in training does not mean that other factors have no significant effect upon the level of skill which the training produces. Whenever different levels of final performance are achieved by a group of individuals who were apparently equal when the training began, it may safely be assumed that individual differences were present at the beginning of the

---

[15] M. K. Bishop, "A Study of Individual Differences in Learning," *Psychological Review*, XXXII (1925), pp. 34–53.

[16] J. Peterson, "The Effects of Practice on Individual Differences," *27th Year Book* (National Society for the Study of Education, Bloomington, Ill., 1928), Part II, p. 212.

[17] Hartmann, *op. cit.*

[18] Morris S. Viteles, *Industrial Psychology* (W. W. Norton & Company, Inc. 1932), p. 107.

training which, although they may have completely escaped notice, nevertheless were instrumental in setting the level of attainment which the several individuals later reached.

An example will make this clear. It has long been standard practice in the hosiery industry to hire girls to be trained as loopers only after rigid scrutiny and interview, careful study of school records and previous employment, general physical and visual examination, and a thorough consideration of all other factors that standard employment procedure has indicated to be of importance. Obviously the girls could not be selected on the basis of their looping ability because, never having seen a looping machine, they have no skill whatever on this job at the time of employment. Except when the labor market is very tight, only those girls are hired who, in the judgment of the personnel manager, are likely to become good loopers with training. But in spite of this extensive examination and investigation at the time of employment, and in spite of the elimination of all those who for obvious reasons are clearly unlikely to become efficient on this particular job, the girls hired do not by any means become equally efficient on the operation. In fact, as Figure 4 shows, after a year of experience some of the girls are more than twice as productive as others. This fact would indicate that, though the girls might have *seemed* to be potentially equal at the time of employment, and indeed probably *were* equal in terms of all factors considered, they were not really equal at all. They differed significantly at the time of employment in one or more of the factors that determine capacity for success upon this job, but—and this is the important point—*the differentiating basic individual differences were not detected by the standard employment procedure.* It should be added that a dexterity test (described on page 130) and certain vision tests (described on page 206) were successful in detecting at the time of employment an appreciable number of girls who were quite satisfactory from the point of view of standard employment procedure but who did not have the basic capacities necessary

for developing a high level of skill in the muscular and visual co-ordination involved in close work.

One should not infer from the foregoing illustration that the industrial psychologist recommends the elimination of standard employment procedure. He fully recognizes that a good personnel man will detect all of the obvious, and many of the subtle, reasons why some applicants should, and others should not, be placed upon a specific job. But experiments clearly show that even the best personnel men cannot determine the muscular co-ordination or dexterity of an applicant from a general interview or from an examination of his or her hands. Many if not most industrial jobs call for a combination of certain basic capacities that cannot be detected at the time of employment without special technique and, often, standardized test procedures. These basic capacities, many of which are measurable at the time of employment, are related to future productivity, accuracy, accident proneness, promotion, versatility, tenure with the company, and many other factors that differentiate a profitable from a nonprofitable employee. One of the major functions of this book will be to show how these basic factors affect the development of an employee and to cite a number of illustrations of the use of selection and placement tests in industry.

2

The Interview and Related Employment Methods

EMPLOYMENT directors recognize many of the facts concerning individual differences discussed in the preceding chapter. If an applicant has training or successful work experience in a certain area, every effort is made to place him on a job where both he and the company will profit from the utilization of his particular experience and skills. Likewise, if a job requires a person with certain physical, mental, or personality characteristics, a man possessing the required characteristics is sought. One of the major functions of the personnel manager is to recognize individual differences in people and in job requirements, and to make every reasonable effort to match men and jobs.

The Interview

As aids in effectively matching men and jobs, several basic methods or tools are available to the employment man. These tools are related to each other and usually are used in a given sequence by the employment department. One sequence often used runs as follows:

1. Application blank
2. Preliminary interview
3. Medical examination
4. Occupational tests
5. Final or employment interview

Not all plants follow this order, nor do all plants use all the

steps mentioned, in their employment procedure. Every plant, however, uses the interview, which is the oldest and still the most widely used of all employment techniques. Even though many studies have been published showing that information obtained through the interview is often inaccurate,[1] that interviewers of presumably equal competence sometimes disagree markedly in their judgments of applicants,[2] and that the interviewer's judgment of an interviewee is often unduly influenced by the latter's physical characteristics,[3] the interview as a basic method still continues to be widely used. This continued use should strongly suggest that the interview, in spite of its limitations, occupies an important place in the majority of employment offices. It would thus seem that, in spite of the possibilities of error inherent in the interview as a technique, it has certain advantages that account for its long-continued use. It is sound industrial relations for every job-seeker to have an opportunity to speak at first hand with his prospective employer (or a representative of his prospective employer) and to be given the personal attention that an application form or employment tests cannot give. Because of these basic considerations, it is advisable to correct as much as possible the attendant weaknesses of the interview, rather than to eliminate it, especially since with it will also go the accompanying opportunity to obtain the good will of the applicant (and, through the applicant, of the community) for the company.

A number of principles have emerged from the many investigations that have been conducted to explore the inter-

---

[1] S. A. Rice, "Contagious Bias in the Interview: A Methodological Note," *American Journal of Sociology*, XXXV (1929), pp. 420–423.

[2] D. A. Laird, *The Psychology of Selecting Employees*, third edition (McGraw-Hill Book Company, 1937), pp. 101–102.

[3] C. Landis, "The Justification of Judgments," *Journal of Personnel Research*, IV (1925), pp. 7–19. S. A. Rice, "Stereotypes," *Journal of Personnel Research*, V (1926), pp. 267–276.

view. Adherence to these principles has been found to improve the value of this tool of employment.

Training the interviewers

Interviewing is like any other skill—it does not come to one "naturally," but must be learned through instruction and practice. And like any other skill, it will not develop properly from "practice" alone if the practice is not properly directed

TABLE 2

SUGGESTIONS TO FOLLOW IN PREPARING TO CONDUCT AN EMPLOYMENT INTERVIEW[4]

1. Decide just what you want to accomplish. What facts do you want to obtain?
2. Know your interviewee. Have you obtained, prior to the interview, all necessary information about him from other sources, such as the application blank?
3. Make appointments. The applicant will judge the company by your reliability in making and keeping the interview appointment.
4. Provide for privacy. Remember, the interview is important to the applicant. He deserves private and courteous consideration.
5. Practice taking the interviewee's point of view. This attitude will help the applicant to overcome any tendency to be unduly nervous or apprehensive.
6. Examine and discount your own prejudices. Do not let your judgment be influenced by prejudice or bias.
7. Gain and deserve the interviewee's confidence. Unless the interviewee has confidence in you, he may not give the information you want.
8. Establish pleasant associations; they will help in establishing confidence.
9. Render your interviewee a real service. Whether or not you hire the applicant, he will appreciate sound counsel and advice.
10. Help the interviewee to feel at ease and ready to talk. Begin with a pleasant topic of mutual interest, allow time for the interviewee to become accustomed to his surroundings, and ask questions that will permit him to show pride in his knowledge.
11. Listen. The importance of this point cannot be too strongly emphasized. Remember that you learn little about the interviewee while *you* are talking.
12. Allow enough time for the interview. If you rush through it, or put pressure on the interviewee, you will not see the latter under normal conditions.
13. Do not dawdle. While the interview should not be rushed, neither should irrelevant topics be allowed to drag it out after the business is over.
14. Keep control of the interview. Channel it back to the business at hand if it goes "off the course."
15. At the close of the interview, watch for additional information or new leads in the casual remarks of the interviewee. Remember that you may learn something about the applicant after you have said "good-by."

---

[4] Adapted from W. V. Bingham and B. V. Moore, *How to Interview*, third edition (Harper and Brothers, 1941), pp. 29–55.

toward the desired goal. Bingham and Moore have listed a number of general suggestions that should be followed in conducting an interview, together with a discussion and explanation of each suggestion. The suggestions of Bingham and Moore that are particularly pertinent to the employment interview are summarized in Table 2; they can be used effectively as a guide in training employment interviewers.

Supplements to the Interview

Most interviewing is markedly improved if use is made of facts obtained from various other sources before the applicant is given the final employment interview. A short, preliminary interview is usually advisable, but then the facts from other sources should be assembled and made available to the interviewer before he conducts the final interview in which a decision on employment is reached.

Among the reasons for obtaining additional facts from other sources is the circumstance that some things about an applicant are readily and accurately revealed in an interview, while other characteristics are revealed in the interview very incompletely, or perhaps not at all. For example, an interviewee quickly shows whether he has a pleasing smile, effective speech and voice, and a general air of confidence. It is not so easy to determine by the use of standard interview technique whether a man can operate a milling machine, read a micrometer, or find the trouble in a defective machine; or whether a girl can type accurately, add a column of figures correctly, or avoid mistakes in spelling. This fact—that the interview is a "natural" for determining certain characteristics but a rather ineffective method of determining others—should be kept in mind when one discusses the "validity of the interview." It would seem to be a mistake to make a single over-all evaluation of the interview as a method. Its value depends on the interviewer (a point already discussed) and also upon the kind of job he is attempting to fill. For example, the very same personal characteristics that enable

an applicant to make an excellent impression on the interviewer will frequently enable him to make an excellent impression on a prospective customer, *if he is hired as a salesman*, owing to the fact that the interview often taps the same personal characteristics that successful job performance as a salesman requires. But the traits that make a good impression in an interview are often far removed from those required for successful job performance as a laboratory technician, an assembler, or a machine operator. Thus a certain type of interview might be highly satisfactory in selecting employees for one job and far from satisfactory in filling another.

To increase the effectiveness of the interview, it is usually advisable to supplement the information obtained directly from the applicant with information obtained from other sources that will be discussed later in this chapter. While true of all interviewing, this fact is particularly important in connection with jobs requiring specific skills, education, or basic training.

Information that should be available to the employment interviewer may be divided into two broad classes—information about the job and information about the man.

Information About the Job

Basic information about the requirements of the job to be filled may be supplied to the employment interviewer from one type of job analysis, namely, job analysis for personnel specifications.

Job analysis for personnel specifications

This type of job analysis results in a specific list of qualifications that a man should have in order to fill the job satisfactorily. The term *job analysis*, however, is used to refer to several other procedures in common use as managerial tools.[5]

---

[5] J. E. Zerga, "Job Analysis; A Resumé and Bibliography," *Journal of Applied Psychology*, XXVII (1943), pp. 249–267.

It is therefore advisable to define briefly what is meant by the major types of job analysis:

1. *Job Analysis for Personnel Specifications.* A determination of the main characteristics (physical, mental, educational, and so forth) that a man must have to fill a particular job satisfactorily. This is the type of job analysis that is used primarily by the employment man. It will be discussed in more detail later in this chapter.

2. *Job Analysis for Training Purposes.* A listing, in order, of the specific steps to be taken in teaching a job to a new employee. This type of job analysis will be discussed in detail in Chapter 8 on Industrial Training.

3. *Job Analysis for Setting of Rates.* This type of job analysis is usually referred to as job evaluation. Its purpose is to set rates or rate ranges for various jobs, so that each is "in line" with the others in terms of their respective requirements and demands. Job evaluation has become a very important managerial tool and will be discussed in some detail in Chapter 11.

4. *Job Analysis for Methods Improvement.* This type is the result of the use of several techniques, primarily motion and time studies, by an industrial engineer charged with the responsibility of analyzing methods of job performance in order to simplify the work layout, eliminate unnecessary motions, and make the job a simpler and more convenient task to perform. This type of job analysis will be discussed in Chapter 12.

Still other meanings are sometimes attached to the phrase "job analysis," as Zerga[6] has pointed out in his resumé, but the four types mentioned above probably illustrate the most commonly encountered meanings of the term. It should be kept in mind that these four types of job analysis are quite different processes—different in method used, in purposes, and in results. It is unfortunate that the single term *job analysis* has come to be used with reference to all four proc-

[6] *Ibid.*

esses, for this indiscriminate use often creates confusion in thinking. One man using the term may be referring to rate setting; another, using the same term, to personnel specifications. A misunderstanding, therefore, may arise that is due to different interpretations of the term rather than to any basic disagreement in intent or purpose.

Of the four meanings of the term *job analysis* mentioned above, the one of primary concern to the employment man is job analysis for personnel specifications. Some jobs require a tall man; others can be done equally well, or perhaps even better, by a short man. Some require considerable physical strength; others do not demand much physical effort. Some require keen vision at specific distances; others are relatively independent of visual skills. Some require arithmetical or computational ability; others require little or no educational background. Before a man can be adequately considered for any specific job, the employment man must have available specific information on what characteristics a man must have in order to fulfill the requirements of that particular job. These personnel specifications can be identified only by a careful study of the job. The first step in this process is to obtain a thorough job description, i.e., a detailed statement of what must be done by the man on the job. The job description should be prepared by someone who is familiar with the actual job and who, in addition, has obtained as much information as possible from the supervisors of the job in question, from operators employed on the job, from time-study engineers, and from safety men. It should be emphasized that the job description is the foundation on which, for the most part, further efforts toward job analysis are based. If the job description is incomplete or inaccurate, any statement of personnel specifications stemming from it will also be incomplete or inaccurate. And if job evaluations (discussed in Chapter 11) are based on faulty job descriptions, no system incorporating them will result in a serviceable and equitable set of rates for the various jobs.

Electrical Repairman[7]

Job Summary

Installs, tests, and services the electric light and power distribution circuits, equipment, and appliances, being responsible at all times for the efficient operation of assigned equipment. Works from blueprints, written specifications, or oral instructions, or independently exercises judgment based on knowledge and experience in determining details of the work. Supervises two helpers.

Work Performed

Duties range from simple wire installation to the installation and maintenance of complicated motor and power equipment. Typical tasks of the job follow:

1. Installs new equipment for building additions and alterations: Measures, cuts, bends, and assembles cable, conduit, and wire, pulling the wires and cable through the lengths of conduit with a fish tape or steel snake, and installs the wiring in floor, ceiling, and wall frames before concrete is poured or before the siding and plastering is put in place; fastens into position such equipment as panel boxes, outlet boxes, switch boxes, and fixtures, performing the tasks as the construction advances; makes final connections to join system with power line, when construction is completed, and closes switches to put new system into operation.
2. Installs and connects motors, fixtures, and appliances: Fastens such equipment as motors, transformers, light fixtures, electric heaters, food mixers, and circuit breakers into position, drilling holes for bolts, shaping and fitting bus bars, and making final connections to the bus bars and wiring system by skinning wire ends and making permanent contacts with solder or patent fasteners; closes main switches and tests installation, making adjustments that are required.
3. Repairs lines, appliances, and equipment: Inspects circuits for defects, using such instruments as ammeters, voltmeters, and test lamps, or locates defects concerning which complaints have been received; remedies faults by performing such tasks as replacing defective wires, cleaning or renewing imperfect contacts, replacing burned-out elements, fuses and light bulbs, repairing or replacing switches, and tightening loose connections.
4. Reads and records meter readings periodically and reports to the management the amount of light and power current consumed.
5. Required to keep records of time consumed at tasks, material used, and appliances installed.

Equipment Used

Ammeter: An instrument used for measuring electric current.

Fish tape (steel snake): A flat, tempered, spring-steel tape or wire used in pulling electric wires and cables through conduit.

[7] Based upon job description of Electrical Repairman 4-97.420, from *Job Descriptions for Industrial Service and Maintenance Jobs* (United States Government Printing Office, 1939), pp. 9–11.

Testing lamp: A light bulb, connected to a pair of wires which are brought into contact with electric terminals or parts of equipment or wires in order to determine the presence of electrical potential by the glow of the bulb.

Voltmeter: An instrument for measuring, in volts, the difference of potential between different points of an electrical circuit.

Working Conditions

Works inside and outside when repairing light and power lines that connect various buildings.

The worker is subject to electric shocks and burns and liable to injury from falls from ladders or scaffolding.

Relation to Other Jobs

Promotion from: ELECTRICIAN APPRENTICE, or from job on which the worker has received equivalent experience.

Promotion to: BUILDING SUPERINTENDENT; CHIEF ELECTRICIAN; MASTER ELECTRICIAN

Specialized Qualifications

Experience on the same job, or as an ELECTRICIAN APPRENTICE, or on a job on which the worker has received equivalent experience, is required.

Must have a knowledge of electrical and building codes and regulations, and of safety laws.

Required to interpret blueprints and written specifications.

Special Information

Worker furnishes tools.

License is required.

It should also be emphasized that the use of abbreviated job descriptions or mere job titles (a job title is simply a very short abbreviation of a description) is entirely inadequate for any type of satisfactory job analysis. If a job title only, or an abbreviated job description, is used, the details of the job must be supplied from the experience (or imagination) of the job analyst. Under such circumstances it is not strange that his efforts do not always result in a satisfactory analysis.

A number of typical and detailed job descriptions have been published in pamphlet form by the United States Department of Labor and the United States Employment Service.[8] The (slightly modified) job description on page 30 illustrates one of the scores of those to be found in these United States Government publications.

[8] These publications are summarized by C. L. Shartle, *Occupational Information* (Prentice-Hall, Inc., 1946), pp. 81–87.

The preparation of job descriptions requires training, experience, and attention to detail. Shartle[9] has given an excellent list of suggestions to be observed in preparing these descriptions:

1. Always be accurate in what is expressed. The facts are what is required.

2. In describing an occupation omit words which are worker attitudes unless it is specified that you are reporting attitudes of workers. Such words include monotonous, uninteresting, interesting, pleasant, distasteful, and the like. Whether a job is monotonous or interesting depends upon the worker, it is not the occupation. One person may find being an executive uninteresting while another may thoroughly enjoy a repetitive task. The writers of occupational information are usually "white collar" workers who may tend to read their own attitudes into the job descriptions.

3. If you are describing an occupation or field of work be sure that you cover it entirely unless you indicate otherwise. That is, do not describe one phase of an occupation and give the impression to the reader that you are covering it all.

4. It is permissible to write occupational information in an interesting fashion but do not try to make the job interesting beyond giving the facts, otherwise persons may be attracted to the occupation who are not really interested in it.

5. Write simply and explain technical terms. Most readers will not be nearly as familiar with the job or occupation as you are.

6. Use a carefully chosen systematic outline of major and minor topics to be covered. This applies whether the material is a short description or a lengthy monograph.

7. Write concisely and give sufficient detail to meet the needs, but never pad and do not skimp on important items.

8. Try out examples of write-ups on the persons who will use the information and obtain their suggestions for improvement.

9. Follow up the use of the descriptive materials and constantly attempt to improve them.

10. Put the date of completion on each description. Revise it as often as changes in jobs and the occupation require. In using source material indicate the date it was gathered.

One method of making information from the job description conveniently available to the employment interviewer is

[9] Shartle, *op. cit.*, p. 77.

to construct for each job a template or job requirement profile that may be placed over the master profile. Figure 13 shows an example of a template of this sort, together with an employment application form. When the form is viewed through the template only the items of information pertinent to the job in question are visible to the employment interviewer.

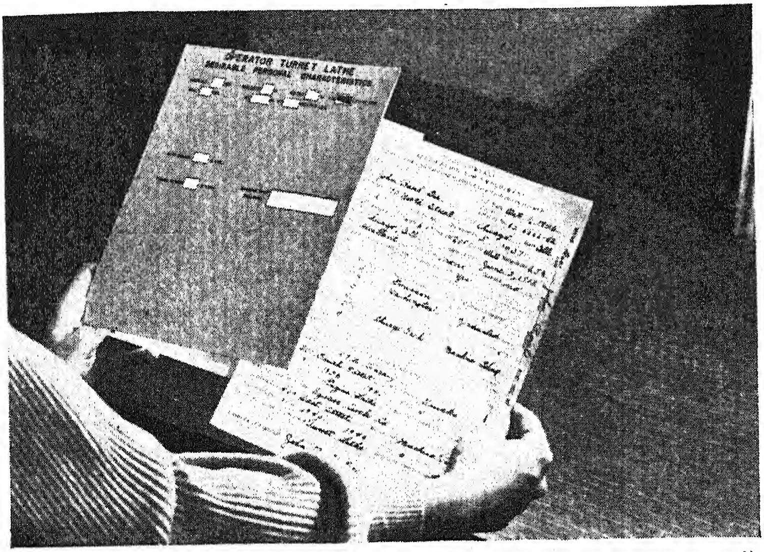


FIG. 13—Template for use by employment interviewer in evaluating application form for turret lathe operator. (Reproduced by courtesy of the American Psychological Association from W. F. Wood, "A New Method of Reading the Employment Questionnaire," *Journal of Applied Psychology*, XXXI [1947], pp. 9–17.)

Statistical analysis of personnel data

Personnel specifications prepared from job descriptions may frequently be supplemented by information obtained from statistical studies to bring out a clear picture of certain physical or educational requirements of a job. To illustrate: Recently a laundry was having trouble on labor turnover among pressers. The job description did not reveal any physical, mental, or educational requirements that could not be met by any normal and physically healthy girl. Yet a

majority of the women hired stayed on the job only a very short time, often leaving without notice.

An analysis of certain personnel data for a group of employees who stayed on the job and another group who left after short employment revealed the facts tabulated in Table 3.

TABLE 3
PERSONNEL DATA ON WOMEN EMPLOYED AS PRESSERS IN A LAUNDRY

	Those Employed More than 8 Months	Those Who Worked Less than 6 Weeks
Average age when hired...	27.2 years	22.2 years
Average height..........	5′ 2.3″	5′ 3.7″
Average weight..........	145 lbs.	125 lbs.
Average years of school....	9.3 years	9.9 years

The obvious interpretation of facts brought out in Table 3 should not be applied indiscriminately to laundries in all localities because labor and personnel conditions may, and frequently do, vary greatly from one locality to another. But for the time and locality in question, it seems rather clear that the employment man should seek as pressers women who, when compared with the average, are older, shorter, heavier, and have had less formal education.

This kind of analysis of personnel characteristics frequently yields clearcut information concerning personnel specifications that is not apparent from a study of the job or even from a very accurate job description. Nor is one limited, as in the illustration given above, to analysis of tenure on the job as the only means of determining the characteristics of the satisfactory employees. Any one of the measures of job success discussed on pages 53–55 may be used in this kind of analysis of personnel data.

Jobs for the disabled

One of the facts brought out by a satisfactory job description is whether the job can be performed by a man with any

given physical handicap. Many plants have prepared check lists of their jobs against the more common types of physical disability, so that the employment interviewer has before him an exact statement of the physical disabilities that can safely and practicably be allowed in a man on a particular job. An example of this sort of job analysis is illustrated in

HANDICAP / POSITION	MILLING AND BROACHING MACH. OPERATORS	LIGHT DRILL PRESS	ENGINE AND TURRET LATHE OPERATOR	BENCH LATHE OPERATOR	HEAT TREAT (INDUCTION WORD	DISPATCHER	CRIB ATTENDANT	BURRING STRAIGHTENING	DELTA BURR MACHINE OPERATOR	BODINE MACHINE OPERATOR
BLIND BOTH EYES										
BLIND ONE EYE							x			
AMPUTATION HAND OR ARM									L.X	
DISABILITY HAND OR ARM						x				
DEAF MUTE										
DEAF										
AMPUTATION ONE OR MORE FINGERS			x					x		
AMPUTATION FOOT OR LEG	x	x					x	x		x
DISABILITY FOOT OR LEG	x			x	x	x	x	x		x
DISABILITY BOTH LEGS										
INJURED BACK OR SPINE										
CARDIAC										
LIGHT WORK RECOMMENDED										

FIG. 14—Job analysis for identifying jobs that can be performed by partially disabled men. (Courtesy of the International Business Machines Corporation.)

Figure 14. This chart and the job analyses on which it is based were prepared by the International Business Machines Company. The chart in its entirety includes information for forty different jobs. In the part of the chart reproduced here for illustrative purposes, the occurrence of an **x** in the column for a particular job indicates that the specified physical handicap may be permitted in a man on that job.

The complete analysis of forty-five jobs revealed that every one of these jobs would permit at least one of the specified physical handicaps, and that many of the jobs could be performed properly by employees with several of the disabilities. For example, seven physical handicaps could be allowed in an assembler, and eight in an operator on a certain part of a coil assembly.

Information of this sort is of value as a means of increasing the potential pool of satisfactory employees. To the employees and to the community it is also of value in identifying suitable work for the many persons who, although their livelihood depends upon work, in many instances have been denied employment because of arbitrary (and sometimes unreasonable) standards of physical fitness.

A similar point of view has been adopted in the extensive work of Hanman,[10] who clearly outlines the procedures to be followed in analyzing jobs for their physical requirements and summarizes the results for a wide variety of jobs.

Information About the Applicant

In addition to information about the job or jobs to be filled, the interviewer should have available as much significant information as possible about the applicant. While certain information can and should be obtained during the interview, other information can often be obtained more conveniently and more accurately by other means.

The application form

The application form, as the name implies, is intended to obtain information about the applicant that will reveal certain facts concerning his employability. The employability of an applicant may, and frequently does, vary from one company to another, from one community to another, and even from time to time in the same company and community. Com-

[10] Bert Hanman, "Matching the Physical Characteristics of Workers and Jobs," *Industrial Medicine*, XIV (1945), pp. 405–430.

pany policy may restrict employment according to certain principles that vary legitimately from one company to another. The application form obtains information that tells the interviewer whether the applicant meets the specifications stipulated by company employment policies.

The application form usually asks for certain facts concerning education and other formal training, previous employers, and work history, and for related information concerning the applicant's previous connections and activities.

The basic assumption underlying the use of an application form as one of the tools of the employing official is that the information obtained is significant in indicating something about the potential value of the applicant as an employee of the company. It would seem logical, therefore, that all questions to appear on the application form be carefully selected, so that only questions eliciting significant information are included. However, even a cursory examination of any sample of application forms in current use selected at random strongly suggests that frequently no such care has been given to the selection of questions included on these forms. Many forms are lengthy documents, and most employment men will admit that little or no use is made of the answers to many of the questions that appear on them. Often, in such instances, the application form consists of a compilation of questions asked on other forms, with the addition of any other questions that may occur to the members of a managerial and industrial relations committee appointed to recommend an application form.

The use of an unwieldy form consisting of a long list of questions (many of which do not yield information really related to the potential job success of the applicant) is open to criticism for several reasons. Applicants are not favorably impressed when they are asked immediately for a considerable amount of personal information about themselves that they feel (often correctly) should be of little concern to the company, inasmuch as it is not related to their qualifications for

the job. A second disadvantage of a long application form, which asks for much unimportant information along with a smaller amount of significant information, is the fact that the interviewer has no means of determining what information is important. He may, therefore, arrive at a decision concerning the applicant that is unduly influenced by the latter's answers to questions that are not important and that should not have been included on the form. A third objection to the long application form arises from the fact that long forms frequently require information about matters extending far into the past. An applicant cannot always remember such details about jobs he held ten years ago as the names of his employers, his wages, and so forth. In order to fill out the form completely, he guesses as best he can, and this guesswork may affect his efforts to be conscientiously honest in filling out those parts of the form that cover matters he can remember more accurately.

The considerations discussed strongly suggest the advisability of using an application form that asks only for information really pertinent to the employability of the applicant. A satisfactory decision as to what questions to include cannot be reached solely by comparing the forms used by other companies nor by accepting blindly the judgments and suggestions of the members of a committee. A satisfactory decision on the content of an application form usually requires a simple but accurate statistical comparison of the later success or failure on the job of employees who give different answers to various questions.

This process requires time, but it is by far the best means of developing an application form that obtains information of real significance to the later job success of the applicant. An example of its use will illustrate the principle. In the development of an application form for prospective life insurance underwriters, a large number of questions were included on the form during its developmental stage. As men were hired and placed on the job (without reference to

the information obtained by the form), the facts concerning which questions asked on the form were pertinent and which were not were revealed. It was found, among other things, that 66 per cent of the married men with dependents succeeded on the job, whereas only 40 per cent of the single men without dependents achieved the same degree of job success.[11] Thus, other things being equal, a married man with children has 66 chances in 100 of succeeding in this type of work, while a single man has only 40 chances in 100 of attaining success.

Many items of information that *seemed* to be of importance did not, in the trial, distinguish the successful employees from the unsuccessful. The information items finally retained on the form were therefore reduced *to those that were proved to be related to job success.* In the case of the information concerning marital status referred to above, applicants were asked whether they were married and had children *not* because someone thought this information was related to job success, *not* because the form could not be considered complete without it, and *not* because it was obtained from one or more forms that other companies had used or were using. This information was included on the form for one and only one reason—namely, that the *sales records showed conclusively that married men with children sold more insurance then single men.*

Another comprehensive series of studies using this method of analysis of personnel data of life insurance salesmen has been directed by A. L. Kurtz[12] and reported in part in manuals[13] of the Life Insurance Sales Research Bureau. This work has resulted in the Aptitude Index, a combination personal-history chart and psychological test, which provides

---

[11] Verne Steward, *The Use and Value of Special Tests in the Selection of Life Underwriters* (Kellaway-Ide-Jones Company, Los Angeles, California, 1934), p. 65.

[12] A. L. Kurtz, *Selection of Agents* (Life Insurance Sales Research Bureau, Hartford, Conn., 1937).

[13] *The Prospective Agents Rating Plan in Use* (Life Insurance Sales Research Bureau, Hartford, Conn., 1938).

a more accurate prediction of potential success in this field of selling than had previously been available. The personal-history chart covers only 10 items. These were chosen after thorough statistical study of 24 personal-history items of 10,111 men employed as full-time agents over a three-year period.

One item *included* in the final personal-history chart deals with the number of different clubs and lodges to which the applicant belongs. One item *not included* in the final chart (although it appears in many application forms) is the number of years of formal education the applicant has had. The reason for *including* the item on clubs and lodges, and *not including* the item on education, may be seen from an examination of Tables 4 and 5. The data summarized in Table 5 show that amount of education *is not* markedly related to success in this particular type of selling.

This conclusion is brought out most clearly by the last columns of the two tables, which show the amount of insurance sold per 100 new employees in each of the categories of formal education and lodge membership. The data show that the number of clubs and lodges to which the applicant belongs is much more markedly related to job success than is the amount of formal education he has had. This kind of analysis of each personal-history item studied resulted in the identification of those particular items that are related to job success for this type of selling. The data on clubs and lodges and the data on education are discussed here to illustrate one method of analysis of personnel data.

These data should not be interpreted as proving (or even as suggesting) that the relative importance of these factors for success in other types of work would be the same as that revealed by these studies. It is entirely possible that the importance of the two factors would be reversed if some other type of work, or even some other type of selling, had been studied.

The examples given above have been discussed in some

detail to emphasize the value of a method that will almost invariably result in a short application form, one that is concerned only with significant information. The discussion is not intended to imply that data on marital status, dependents, clubs and lodges, or education are certain, or even

TABLE 4*

RELATION BETWEEN NUMBER OF ORGANIZATIONS (CLUBS, LODGES, ETC.) THE AGENT BELONGS TO AT TIME OF EMPLOYMENT AND HIS SUBSEQUENT SUCCESS AS A LIFE INSURANCE AGENT

Number of organizations	*Number of agents hired*	*Per cent of agents surviving two years*	*Average sales of two-year survivors for two years*	*Per cent surviving for two years* times *sales for the two-year period*
0	373	11	$112,000	$1,232,000
1	1404	12	91,000	1,092,000
2	651	16	91,000	1,456,000
3	260	26	103,000	2,678,000
4 or more	152	24	166,000	3,984,000

* The data in this table and in Table 5 are taken from material furnished in a personal communication from Dr. A. L. Kurtz.

TABLE 5

RELATION BETWEEN AMOUNT OF FORMAL EDUCATION AND SUBSEQUENT SUCCESS OF NEW LIFE INSURANCE AGENTS

Number years schooling completed	*Number of agents hired*	*Per cent of agents surviving two years*	*Average sales of two-year survivors for two years*	*Per cent surviving for two years* times *sales for the two-year period*
8	670	13	$129,000	$1,664,000
9–11	379	13	89,000	1,157,000
12	3,122	11	123,000	1,353,000
13	372	13	104,000	1,352,000
14	648	13	94,000	1,222,000
15	270	14	143,000	2,002,000
16	3,567	15	122,000	1,830,000
17–22	108	16	135,000	2,160,000

likely, to be of significance in filling other jobs. But an analysis of the type described will show whether such information is important, and also what other types of information are and are not important.

In addition to information that is known to be of value to the employment interviewer, certain additional information is always needed by the company for its personnel

records, including such items as social security number, address, and telephone number. The application form should of course provide for the recording of personnel information of this sort.

Work history

Information concerning an applicant's work history is often of value to the employment interviewer, largely because, other things being equal, what a man has done in the past is a rather good indication of what he will do in the future. Many of the most accurate methods of predicting human behavior are based on this principle.

Work-history information is often obtained most conveniently from the application form. The same rules to be followed in deciding *what* to ask that have been discussed in the preceding section should therefore also be applied in evaluating the significance of work-history information. An example of the results of this type of analysis applied to the latter has been published by Steward[14] and is summarized, in part, in Table 6.

The results in Table 6 show that in selecting men for this type of selling, men with prior work experience in such fields as real estate selling, sales management, and other occupations appearing near the top of the list, have a much greater chance of success on the job than do men with previous experience in fields appearing near the bottom of the list. Previous experience in fields indicated near the center of the list is not related either positively or negatively to success in selling life insurance. The use of information of the sort summarized in Table 6 is of considerable value to an interviewer in evaluating an applicant's work history in relation to possible success on the new job.

---

[14] Verne Steward, *Selection of Sales Personnel Reports*, Bulletin 1, January, 1947, p. 2. (Distributed by Verne Steward & Associates, Box 226, South Gate, California.)

TABLE 6

RELATION OF TYPE OF PRIOR EXPERIENCE TO SUCCESS IN LIFE INSURANCE SELLING

Previous Occupation	r	Degree of Relationship
Real est. salesman	+.71	
Sales manager	+.67	High overlapping
Advertising mgr	+.35	(Favorable occupations.
Lawyer	+.22	Many qualify for life
Social sci. teacher	+.19	insurance)
Banker	+.18	
Y.M.C.A. secretary	+.18	
Pub. utility sales	+.15	Med. overlapping
City school supt	+.13	
Personnel mgr	+.10	
President mfg. co.	+.09	
Office worker	+.06	
C.P.A.	+.04	Low overlapping
Purchasing agent	+.02	
Author-journalist	+.01	
Y.M.C.A. phys. dir	+.01	
Minister	+.01	
Accountant	.00	No relationship
Musician	−.05	
Artist	−.07	
Osteopath	−.10	Slight aversion
Policeman	−.11	
Production mgr	−.14	
Doctor	−.26	Medium aversion
Architect	−.31	
Psychologist	−.32	
Dentist	−.34	
Forest service	−.36	
Math.-sci. teacher	−.41	
Printer	−.45	
Farmer	−.50	(Unfavorable occupations.
Carpenter	−.53	Few qualify for life insur-
Mathematician	−.58	ance)
Eng. (civil, mech.)	−.61	
Physicist	−.69	
Chemist	−.71	Great aversion

Letters of recommendation

In employing men or women for many kinds of jobs, particularly higher-level jobs, great reliance is often placed upon letters of recommendation. Theoretically it would seem that a letter about an applicant written by a former employer, or by some responsible person who has been in close contact with the applicant, would provide a highly satisfac-

tory method of obtaining information that is accurate and comprehensive. To accomplish this result, however, the writer of the letter must be both able and willing to put into writing his full and frank opinion. Such ability and willingness are not always present, and there are even some circumstances under which writers of letters about applicants are inclined to give something less than a comprehensive and accurate picture of an applicant. This fact is indicated by the results of a questionnaire survey of 100 sales managers, partially summarized in Table 7.

TABLE 7

RESPONSES OF 100 SALES MANAGERS TO QUESTIONS CONCERNING THEIR PRACTICE IN WRITING LETTERS OF RECOMMENDATION[15]

Questions Asked the Sales Managers	*% Yes*	*% No*
In replying to a letter of inquiry from a prospective employer regarding the fitness of one of your former employees,		
Do you tell only the good things you can say about the man?	77	23
Do you always give the employee the benefits of any doubts?	85	15
Do you point out the man's failings and weaknesses as well as his strong points?	39	61

The unwillingness of many to write a full and frank statement is probably due in part to a natural reticence to put into writing comments that necessarily become a matter of record, and partly to the fact that the employer of an undesirable man may write for him a favorable letter in order to facilitate his moving to another location. This subterfuge, however, is very poor personnel practice and is not encountered frequently today.

Letters of recommendation are not commonly used in plant and shop employment. The hazards of this tool, therefore, need be guarded against only occasionally by the plant employment interviewer, but he should be aware of the fact that such letters are not always complete and exact, and, when such letters are offered for his consideration, he should make every attempt to verify the information from as many previous employers as possible.

[15] By a New York Sales Manager, "To Whom it May Concern," *Sales Management* (October 1923), p. 9.

Personnel tests

Personnel tests, when given and interpreted properly, constitute one of the most effective aids available to the employment interviewer. Such tests very often quickly and accurately reveal information about an applicant's skills, abilities, aptitudes, and interests that is very difficult to obtain by any other means. This fact does not mean that personnel tests constitute the final answer to the employment problem. Certain characteristics of an applicant that are important in determining potential success or failure on the job are not entirely revealed by any personnel tests that have been developed up to this time. But as a supplement to other employment techniques, and as an aid to the man conducting the employment interview, personnel tests have been proved in scores of practical situations to be many times worth their cost.

The application of psychological methods in the development and standardization of personnel tests constitutes one of the most significant contributions of industrial psychology. The importance of this field, the great amount of attention that psychologists have given to it, the extensive research on tests that has been published in psychological and personnel journals, and—perhaps, above all—the fact that a proper use and interpretation of personnel tests requires careful study, justify the rather extensive coverage of this subject contained in Chapters 3 through 7.

3

General Principles of Employee Testing

THE nature and significance of individual differences among industrial employees were indicated in the first chapter. Since every personnel man recognizes the existence of such differences, it is quite unnecessary for the industrial psychologist to re-emphasize the importance of proper employee placement. The personnel procedures in use in the modern employment office have been installed primarily for the purpose of making a careful and adequate selection and placement of employees. The psychologist's function, rather, is to bring to the attention of the personnel manager those aspects of the theory and application of modern psychology that have been found helpful in the performance of this difficult task.

General Concepts

Testing supplemental to present employment procedures

Many industrialists are skeptical about the use of psychological tests in employment because they do not wish to risk a change from present methods which are known to be reasonably satisfactory to new methods that have not been thoroughly tried in actual practice. Their skepticism is clearly justified. When psychological methods are used in an employment office they should always be looked upon as supplementing—not replacing—other methods that are in use. No psychologist who has thought at all about the problems of

modern industry would seriously suggest that present employment procedures should be eliminated. Indeed, many of the advances that modern industry has made during the past quarter of a century can be traced directly or indirectly to the procedures now in use in every modern industrial employment office. But the fact that these methods, though excellent in many respects, are still not perfect is proved by the marked individual differences among employees which any study of differential production will reveal. A considerable amount of research, both in industry and in the laboratory, has shown that still further improvements in employment methods can be attained when psychological tests and methods are used as supplements to other employment procedures. Let us keep in mind this word "supplement." Psychological tests are among the tools that are necessary for the most effective selection and placement of employees.

Selection or placement as the function of tests

Employee tests have often been considered primarily as devices to aid in the selection of employees. This viewpoint is followed by the assumptions that (1) when tests are used systematically, many applicants will be rejected—that is, not employed at all—and that (2) unless there are significantly more applicants than there are jobs to be filled, the testing program loses its effectiveness.

It is unfortunate that the selective rather than the placement features of employee tests have been given the greater emphasis in most discussions of this subject. Perhaps this is due to the fact that many industrial testing programs were inaugurated during the depression years of 1930 to 1935, at which time there was an unusual abundance of applicants for nearly every job. But in a period of emergency production there is usually not an abundance of applicants. Indeed, many industries find it necessary to relax all employment standards, including not only psychological test results (in industries where these have been established) but also stand-

ards dealing with medical and physical requirements, age, and marital status. When a testing program has been installed on the basic premise that it is a *selection* program, the fact that it is also an excellent *placement* program is often forgotten. Yet the value of a testing program in placing employees during a period of rapid plant expansion is fully as great as its value in selecting employees during more static periods. One of the most effective uses of a testing program that the writer has observed was in a war industrial plant that was hiring, as the medical director stated, "anyone healthy enough to walk into the plant." In this plant the sole purpose of the tests was *placement* of the many new employees on the specific jobs to which each was best adapted. The value of tests is by no means limited to situations in which there is an abundance of job applicants.

Psychological methods are not infallible

A further point which the advocate of testing procedures should make clear is his recognition of the fact that psychological tests are not infallible; that they sometimes give results that are not a true indication of the potential job success of the applicant. Any new procedure, whether in employment, production, advertising, or the like, should be evaluated in terms of not whether it achieves perfection, but whether it results in some improvement over methods that have preceded it. Thus, if the labor turnover in a given department has been 25 per cent per year among employees placed by previous methods; if it is found that new employees placed by psychological tests show a turnover of only 20 per cent per year (all other factors which indicate a desirable employee remaining constant); and if the expense of administering the testing program is less than the amount saved by the reduction in labor turnover, the testing program would ordinarily be considered a sound investment even though it did not achieve perfection in reducing labor turnover to zero. Evaluation of a testing program should be made in terms of a statistical

comparison between the employment situation with the tests and the employment situation without the tests. This statistical comparison should include such factors as the productivity of employees, labor turnover, accuracy and safety, ease of shifting employees to different types of work, ease of learning the job, and any other factors that may be indicative of a desirable employee. It seems only fair to evaluate a testing program in terms of averages rather than in terms of specific cases. This point is emphasized because unless one is careful he is likely to allow one serious failure in a testing program to outweigh the less spectacular though much sounder averages that really indicate the value of the method.

Types of tests

Tests that have been and are being used for employee placement may be classified in several different ways. They may be *group* or *individual* tests. The group variety may be given to almost any number of persons simultaneously, the only limitation on the number usually being seating and writing facilities and provision for adequate hearing of the instructions given by the group examiner before the test is begun. Examples of group tests are the Purdue Achievement Tests (see page 155), the Adaptability Test (see page 85), the Purdue Mechanical Adaptability Test (see page 107), the Otis Self Administering Tests of Mental Ability (see page 83), the Wonderlic Personnel Test (see page 84), and the Bennett Test of Mechanical Comprehension (see page 101). Individual tests, on the other hand, are given to one person at a time and usually call for the undivided, or nearly undivided, attention of the examiner while the test is being administered. The phrase "nearly undivided attention" is used because in certain cases, as with the Purdue Pegboard Test of Manual Dexterity (see page 126), it is possible for an attentive examiner to test several persons simultaneously if the necessary sets of equipment are available.

Another classification of tests may be made according to whether they are *instrumental* or *paper and pencil.* The former, as the name implies, make use of instruments and ordinarily are individual in nature. The latter make use of written responses and ordinarily are group tests.

A very important division of tests may be made according to whether they measure *aptitude* or *achievement.* Aptitude tests measure whether an individual has the capacity or latent ability to learn a given job if he is given adequate training. Such tests are most useful when the majority of applicants for a certain job have had no experience on the job and when a relatively long period of training is required before their aptitude or lack of aptitude for the job will be apparent on the job itself. A good example of a job for which aptitude tests are particularly adapted is looping in a hosiery mill. This is a job which requires a year of experience, on the average, for complete mastery. New employees are usually hired directly out of high school with no previous experience on the operation of a looping machine or anything even remotely resembling it. Aptitude tests that have been found to be effective for this job deal with such factors as finger dexterity (see page 130) and certain visual characteristics (see page 207). Interestingly enough, the tests that "came through"[1] (see page 130) in no way resemble the actual looping operation; rather they measure certain basic capacities that are necessary for an efficient performance on that job.

Achievement tests, on the other hand, measure how well the individual can do the job or what he knows about it at the time he is tested. Achievement tests are of greatest value when many of the applicants for a job have had, or claim to have had, experience on the same or a similar job in some other organization. For example, a standardized test of knowledge pertaining to a machine shop will quickly reveal

[1] The phrase "come through" means that the test scores are found significantly related to success on the job.

how much the applicant really knows about the machinery and operations in use in a standard shop. It is true that he may evidence considerable knowledge on such a test and still be a poor machinist because, for example, he lacks manual skill. But experience shows that if he reveals little or no knowledge on such a test he is exceedingly unlikely to be a good machinist. This type of achievement test, therefore, is of definite value in eliminating from consideration those who lack the basic information necessary for satisfactory performance on the job. Achievement tests, such as the punch press test described on page 150, may also be instrumental in nature.

Although the foregoing discussion may seem to imply that a given test may always be definitely identified as an aptitude test or an achievement test, this implication is not always valid. There are certain tests that, when they are used in certain ways, may most properly be considered achievement tests, but that, when they are used in other ways, are more properly classified as aptitude tests. For example, a standard typing test, such as the Thurstone Examination in Typing (see page 154), would definitely be an achievement test if it were used to select applicants for a typing job. This same test, however, might quite properly be considered an aptitude test if it were used as one element of a test battery designed to select teletype trainees, for in this latter case the typing skill the test measures is basic to the successful accomplishment of the trainee. It should be kept in mind that the basic distinction between aptitude tests and achievement tests must be made in terms of the *purpose* the test is intended to serve—not in terms of the content of the test itself.

Definition of terms

The term *capacity* refers to potential or latent ability. One might have high capacity for a certain skill although he may have none of the skill at the time of testing. Thus, one might have exactly the combination of keen vision at close

distances, finger dexterity, and mechanical ability necessary to become a watchmaker or repairer, but, if he has had no training in watchmaking, he might know nothing about the mechanism of a watch. It is always advisable to know how much capacity one has for a certain job before training is given on that job. Ordinarily those with capacity for the the job will learn quickly and efficiently; those without capacity for it will not.

The term *ability* refers to developed capacity, or actual knowledge or skill already developed. One with little capacity might, if he has been given a great deal of training, show more ability at the time of testing than one with high capacity who has not been trained; but if the latter is trained he will soon excel the former. Therefore, from a long-time viewpoint, it is more important to know the capacity of applicants before they are placed than to know their ability. For purposes of immediate addition of personnel who will get out some production, the opposite situation is of course true.

A relationship may be inferred between types of tests and the definitions of capacity and ability given above. Aptitude tests ordinarily measure capacity; achievement tests measure ability.

Testing the Tests—the Experimental Approach

It is of utmost importance in any testing program to use only tests that have themselves been tested, or to make provision for testing the tests before finally accepting them as valid devices for employee placement. Two general methods may be followed in testing the tests. One of these methods consists of measuring present employees and correlating the test results against whatever criteria of desirability in the employees may be available. The other method consists of testing new employees at the time of hiring, filing the test results, and later determining the relationships between the test results obtained at the time of hiring and the success of the employees on the job after they have been on the pay roll

for a period of time. Each of these methods of testing the tests has advantages and disadvantages. A long-time testing program should make use of both methods. Only by so doing can an employment manager hope to obtain the maximum benefits from a testing program.

The criterion

Both of the two basic methods of "testing the tests" require the use of a suitable criterion of employee desirability. It is therefore advisable to discuss what is meant by a criterion and to consider some of the criteria that have been used in personnel testing.

A criterion is a measurement (*entirely apart from any test results*) of how satisfactory an employee is on the job. Unless and until one has criterion data on a group of employees, he has no yardstick, no standard, against which to check the test scores. Many different criteria have been used in testing personnel tests. There is probably no single criterion that is always to be preferred to any other. The choice of a criterion, or combination of criteria, must be made so as to indicate what management considers to be the most desirable type of employee. Thus the choice of a criterion is necessarily a judgment made by management, just as market predictions, distribution costs, and other similar factors must and should rest on managerial decisions.

Among the criteria that have been used most frequently are the following:

1. *Production.* If the other things listed below are equal, the production achieved by an employee is usually considered indicative of his value to the company. Unfortunately, however, there are countless hourly-paid jobs in industry for which no individual production records are available. When individual production records do not exist, production obviously cannot be used as a criterion.

2. *Quality of Work.* If two employees are equal in production (or if they are both paid a day rate and their

actual production is not measured), the one who has the *least* rejects, or who wastes the *least* material, is presumably the better employee. The quality of work performed may therefore at times be used as a criterion of employee desirability even when individual production records are not available. Sometimes when there are individual production records for each employee, as well as quality records (in the form of number of rejects or amount of wasted material), it is desirable to establish a criterion by combining production and quality records to form a single standard. When this is done, the two records combined should be weighted according to management's judgment of their relative importance, using the method described in Appendix A, page 502, in order to be sure the weights decided upon are actually given.

3. A third criterion used by some managements is *learning time, or its equivalent, training cost.* Other things being equal, if one new operator learns the job at a cost to the company of $50, while another requires training that costs $150 before he reaches the same level of production and quality, the first operator is the more desirable employee.

Because of the minimum wage provisions of the Fair Labor Standards Act, it is often of great importance to economy in production to obtain trainees who are able to "earn" the minimum rate in the shortest possible time, because by law this minimum rate must be paid to the employee from the beginning of his employment.

4. *Tenure on the Job.* Suppose two new employees learn a job with the same speed and at the end of a period of time, say six months, have reached equal levels of production and quality. If one of these employees now quits while the other stays on the job several months, or perhaps years, longer, it is clear that the employee who stayed longer on the job is the better investment from the company's viewpoint. In the case of the short-term employee, it is often possible that more money is invested in training him than can possibly be realized on his low production while in training. For this

reason, tenure on the job is often used as a criterion of employee desirability.

5. *Absenteeism.* The percentage of working days that the employee actually reports for work has proved at times a very usable criterion. If one employee is present 97 to 98 per cent of the time and another present only 80 to 85 per cent of the time (other characteristics of the two employees being equal), the former would usually be considered the more desirable employee.

6. *Accidents.* When personnel testing programs are installed either by the safety department, the medical department, or the personnel department working in collaboration with either of the other two, one of the major objectives is to select new employees who will be relatively free from accidents. Under these circumstances, low accident experience is frequently taken as a criterion of a desirable employee.

7. *Promotions.* When a hiring program has as one of its major objectives the selection of men or women who will rapidly develop the ability to move into higher-level jobs, the speed of such upgrading may be the most satisfactory criterion to use.

8. *Job Sample.* When data on production and/or quality cannot be readily obtained, it is sometimes possible to obtain measurements of employee job preformance by a specially set-up job sample. The use of this criterion is illustrated in Chapter 9.

9. *Merit Ratings.* One of the most widely used of all criteria is merit rating, the systematic rating of an employee's ability by his immediate supervisor or some other representative of management who is thoroughly familiar with that employee's work. Merit rating, if properly used, gives management a tool that is usable not only in fashioning a criterion, but also in providing aid in problems of transfer, promotion, and employee improvement. Chapter 10 will be devoted entirely to this subject.

The foregoing discussion of criteria will serve to focus attention upon some of the more commonly used methods of

measuring the value of an employee to the company. Any effort to "test the tests," as described in the following section, must rest upon the acceptance of one or more criteria as a measure of job performance.[2]

Testing the tests on present employees

This method involves testing a group of employees and determining the relationship between test results and employee efficiency. When a plant begins, for the first time, a project of test validation that involves testing a large group of present employees, the question of the reaction of the employees naturally arises. Will they become anxious or unduly excited? Is there a possibility of the testing program causing worker unrest? In this book will be summarized research work and test results that have been obtained by testing hundreds of employees in numerous plants. Many of these plants were strongly unionized. In the union plants, several labor organizations have been represented. In no case has any trouble arisen either with employees or union representatives. Wherever the management has signed a union contract, representatives of the union are contacted and the project explained to them before actual testing is begun. Usually a mimeographed slip reading somewhat as follows is given to each employee:

> The Personnel Department is conducting a series of experiments. You have our assurance that this testing is being done to "test the tests" and that the results will *not* be used, now or later, in any way that will affect your standing with the company.
>
> PERSONNEL DIRECTOR

It should be emphasized that in using this method of testing the test, management should be utterly sincere in its plan to use the results *only* to test the tests. Under no circumstances should the test results be used as a basis for transfer or layoff, or for any other purpose detrimental to the

[2] An excellent discussion of criteria by J. L. Otis has been published in Chapter V of W. H. Stead, C. L. Shartle, *et al. Occupational Counseling Techniques* (New York: American Book Company, 1940).

employee or which he may *believe* to be detrimental. After a testing program has been thoroughly tried out and installed, many uses for the test scores may be made, uses that will be advantageous to both employees and management; but in the "testing the tests" stage of the program, management should adhere scrupulously to its agreement not to use the results for any other purpose than the one stated.

After the test results have been obtained, the results may be analyzed in several different ways. One method is to divide the employees into two groups according to whether they score above or below average on the test in question, and compute the average job performance criterion of both the "high-testing" group and the "low-testing" group. The criterion used to determine these scores is usually selected in conference with management (or at times with management and the union) from the list discussed in the preceding section. If a significant difference between the scores of these two groups is shown, it may safely be assumed that the test is measuring something of importance on the job. Under these conditions it is usually desirable for the employment manager to know whether an applicant is in the high- or the low-testing group before the applicant is placed.

Another, and even more effective, method of determining the relationship between the test results and employee efficiency is to determine the coefficient of correlation between the test scores and the efficiency of the employees. This method has numerous advantages over simply dividing the employees into a high-testing and a low-testing group. In the first place, it gives a more accurate indication of the *amount* of the relationship between test scores and efficiency. In the second place, it enables the employment manager more effectively to take advantage of the all-important selection ratio (see page 66) in using the test. In the third place, it makes possible the computation of the relative importance of several tests in an employment battery so that the tests may be "weighted" according to their importance. Finally, the

use of the correlational method makes possible the use of partial and multiple correlation to eliminate statistically whatever effect such factors as experience on the job or age may have had in determining the correlation between test scores and job performance.

From the logical viewpoint, one objection to the "present employee" method of "testing the test" is that the test may be measuring something that is improved significantly by experience on the job. In other words, the test may be an achievement test rather than an aptitude test. Consider again the operation of looping in a hosiery mill. If a test is proposed which imitates the operation of a looping machine, it is almost certain that present employees will divide, in their ability to score well on this test, in much the same way that they have already divided in their ability to do the actual job. This simply means that the test, being in itself a miniature of the job, will divide the employees in much the same way that they have already been divided on the basis of the job itself. Such a test would have little value for placement of new employees unless the surplus labor supply included a large number of employees who claimed to have had looping experience in some other hosiery mill. Under ordinary circumstances, however, persons employed for this job are freshly out of school. They have had no experience in looping, and their capacity for the job would not be measured effectively by a test of this type. Since none of them can loop at the time of employment, all of them, even those who are potentially the best employees if given the training, would score zero or close to zero on any test that calls for ability to loop at the time of employment.

In testing tests according to this principle it is therefore necessary to be sure that the tests, in addition to dividing the employees according to their ability on the job, do not show a significant correlation with experience on the job. In other words, it is necessary to be sure that the employees who score high on the test are not scoring high simply because they have

had an opportunity to learn the skill being tested. Whether this is the case may be determined by correlating the test scores with experience on the job. If the test is to be used later for employment and/or placement it must satisfy two requirements. First, it must show a positive correlation with ability on the job. Second, the scores on it must not be appreciably related to experience on the job. When these two conditions are satisfied, it may be assumed that the test is *not* measuring something that is improved markedly by experience on the job, but that it *is* measuring something that is necessary for adequate performance on the job. This reasoning applies only to *aptitude* tests, not to *achievement* tests. When tests of the latter type are used, a significant correlation between test scores and experience is to be expected and in no way reduces the value of the tests.

If the test scores show *some* correlation with experience, the net relationship between test scores and job performance, after the effect of experience has been eliminated, can be determined by partial correlation. The procedure for computing partial correlations may be found in any standard textbook of statistics.[3]

The foregoing discussion applies only in cases where the applicants have not had experience on the job for which they are being hired. If they have had experience, then an achievement test that is similar to the job or even a *miniature* (that is, the job itself in standardized test form) of the job may identify the best employees even more satisfactorily than an aptitude test.

Testing the test on new employees—the "follow-up" method

A second and, in general, a more effective method of determining the value of a test, is to test the test on new employees. This method consists of giving the test, at the

[3] Henry E. Garrett, *Statistics in Psychology and Education*, third edition (New York: Longmans, Green, and Company, 1947).

time of employment, to a large number of employees whose placement has already been decided upon by ordinary methods of selection and placement. These employees are hired just as they have always been hired, but before they have been put upon the job or, in fact, before they have been told that they have been selected for employment, they are required to take a battery of tests. These tests at this time have no effect whatever upon whether the applicants will be hired. As already mentioned, it has been decided to hire those who are to be tested before the tests are given, but these employees are not informed of this decision. When this method of testing the test is followed, the test results are filed and, for all practical purposes, forgotten, until the employees have had an opportunity to show whether or not they are successful on the job. When sufficient time has elapsed so that it is known which of the employees have been successful, which average, and which unsuccessful, on the job, the test results are taken from the file and the statistician goes to work. His job is to determine the relationship between the test results obtained at the time of employment and the later success or failure of the employees on the job. The statistician makes use of one or more criteria of the type discussed on page 53 in determining the relationship between test scores and job success. The statistician may consider the wage earned by the employees, the amount of production, the proportion of working days the employee has been on the job, freedom from accidents, rapidity of learning the new job, ratings by supervisors on quantity and quality of work, tenure of the employee before layoff or resignation, or any other of several indications of success in an employee that a given company may feel to be important in its particular situation.

This method requires more time than testing the tests with present employees. Frequently, it does not result in any proved tests until the program has been under way for several months or, in some cases, several years. Should the first

battery of tests selected for tryout prove entirely unsatisfactory, it is necessary to start the whole procedure over again. This involves the loss of a great deal of time.

Fortunately, in practice the employment manager need not decide upon one or the other of the methods discussed and then limit himself exclusively to that particular method. He may proceed with both methods simultaneously. When this is done it is possible to produce a working battery of tests almost at once (by using the present-employee method of testing the tests) and to obtain still more evidence of the value of these tests as new employees are tested at the time of placement. This statement may seem to contradict itself because it may be argued that if a test is used for placement there will be no variation in the scores on this test among the new employees placed on any given job and, therefore, it will be impossible to correlate the scores with later success on the job. Actually, even though test scores are used as a basis of employment or placement, there will ordinarily be a sufficient spread of scores on the test, even among employees who are hired for a specific job, to make possible the subsequent statistical evaluation of the test in terms of the follow-up method. Though the variability of the group of employees placed on any given job will be smaller than that of all the applicants, nevertheless the variability of all the applicants and the variability of the applicants who are placed on the job will be known and it will therefore be possible to determine on the basis of the follow-up method just how much the test in question increases the efficiency of the placement procedure for the particular job.

The point to be emphasized throughout this discussion is that no one—whether he is an employment manager, a psychologist, or anyone else—can predict with certainty which tests will be desirable tests for placement on any particular job. Often a rather accurate estimate can be made; but it is always necessary to check this estimate against correlations between the actual performance of the employees on the job

and the scores which they made on the tests at the time of placement.

Factors Determining the Functional Value of Selection Tests

Several factors of both a practical and theoretical nature determine the functional value of a testing program in an industrial plant.

1. When many applicants are being hired for a variety of jobs

Tests furnish a highly effective means of placing employees in the particular jobs for which each is best adapted. Thus, in the case of an assembly line, some jobs may require keen "near-point" vision, others manual strength, and still others good color vision. Sometimes still different factors or a combination of factors are involved. Even when there is no thought of using tests to "pick and choose" from among applicants, the placement value of tests during a period of rapid plant expansion is very great. Indeed, it is during such a period, when it is necessary to hire and place a large number of employees about whom relatively little is known, that tests are of greatest value to the personnel manager.

2. Validity and reliability

Further factors determining the value of tests are their *validity* and their *reliability*. A test is valid if it measures what it is supposed to measure. It is reliable if it measures *something* accurately and consistently. Much of the preceding discussion under the heading "testing the tests" will be recognized as a means of determining the validity of industrial personnel tests. The industrial psychologist is in a favorable position with respect to determining the validity of his measuring instruments. He recognizes the need for a satisfactory criterion and is able to examine available plant records with this need clearly in view. When no satisfactory criterion is readily available, the psychologist often is able to assist management in establishing a criterion from existing records

or from new personnel or production data. While criteria are often established for the primary purpose of "testing the tests," once established, they often are found of real value for several other purposes such as the analysis of personnel data in relation to costs of production.

The validity of a psychological test is usually expressed as the coefficient of correlation between the scores on the test in question and some outside criterion—that is, an indication or measurement of what the test is supposed to measure. It has already been mentioned (see page 53) that these criteria include wages earned, quality of work, freedom from accidents, merit ratings, or other factors that the company accepts as indicative of a desirable employee. In using any test it is necessary to know not only the validity of the test, but also just what criterion has been used in determining that validity.

Tests used in industrial placement must also be reliable; that is, they must measure consistently whatever they measure. A test is reliable if it consistently gives the same score to an employee when he is retested. The retesting may be done with the same test used originally, if practice or memory does not markedly affect the test score. If practice or memory factors significantly affect the score, however, retesting should make use of a duplicate form of the test containing items similar in nature but different in actual content.

Since no test, no matter how perfect, will give *exactly* the same score on a first and second trial to all persons tested, a method of measuring reliability is used which expresses the degree of reliability or the *extent* to which the test gives the same score to an employee on repeated testing. The reliability of a test is expressed as a coefficient of correlation, which may be interpreted as the correlation between scores obtained from repeated testing with the same or duplicate forms of the test. When duplicate forms are not available but there is reason to believe that the memory factor will influence the scores, a method known as "chance-halves" or

"odds vs. evens" is often used in determining the reliability of a test. This method involves correlating certain parts of a test against the remainder. The test-retest and chance-halves methods of determining reliability do not give identical results. Technical descriptions of the research work involved in standardizing a test state not only the reliability of the test but what method of determining the reliability was used.

A test may have high reliability and yet be quite worthless for any given employment situation. A test might measure height, weight, or even general intelligence with high reliability and yet show little or no correlation between test scores obtained and success of the employees on the job. In other words, the test might have high *reliability* but low *validity*, or even no validity at all. But if a test has a *low* reliability, it is not likely to have a satisfactory validity. This is simply another way of saying that no test is likely to have a higher correlation with anything than that test will have with itself. Thus a test which on repeated testing gives scores that correlate with the first testing only .45 is not likely, except by chance, to relate higher with production or any other criterion of employee desirability.[4] It is primarily for this reason that the psychologist using tests in industry insists that they be reasonably reliable before he makes any attempt to determine their validity for particular situations.

A question that is often raised when the validity or reliability of tests is under discussion is: How reliable and valid must a test be in order to be worthwhile? This is both a

[4] Theoretically, the coefficient of validity of a test might reach a value equal to the square root of the coefficient of reliability, although this ceiling is seldom if ever reached in actual situations. Thus a test with a reliability coefficient of .64 might conceivably have a validity coefficient of .80. However, even though the validity might theoretically exceed the reliability, the square root of the latter still determines the ceiling of the former. Therefore, other things equal, it is desirable to have the reliability of a test as high as possible before an attempt is made to determine its validity for a particular job. For a more extended discussion of this topic, see E. F. Lindquist, *A First Course in Statistics*, (Boston: Houghton, Mifflin & Company, 1942), pp. 223 ff.

reasonable and a natural question when one considers that our measurements of both reliability and validity are usually expressed in the form of a coefficient of correlation which makes possible a variation in the degree of reliability or validity all the way from zero (or even a minus quantity, if we wish to be theoretically exact) to 1.00. If we remember that reliability in a test is necessary because it limits the validity, we may phrase the question more simply by asking: How high must the validity coefficient of a test be for the test to be worthwhile?

The answer to this question depends upon the use that is being made of the test. The user of tests is nearly always interested in one of two objectives, but is seldom interested in both at the same time. Either he is interested in making a careful and accurate aptitude analysis *of each person tested*, which is to be used for individual prediction or vocational guidance, or he is interested in selecting from a large group of individuals a smaller group which, *on the average*, will excel the larger group in some particular respect. In individual consulting work, which deals with vocational aptitude and guidance, the psychologist is interested in the first-named objective. His work will stand or fall on the accuracy of his predictions for individual clients. He therefore has little use for aptitude tests that do not have a validity sufficiently high to justify their use in individual prediction. The exact value which the validity of a test should have to meet this requirement is not completely agreed upon by all students of the subject. It is uniformly agreed, however, that the higher the validity of the test the better and *that there is no substitute for validity for individual prediction.*

On the other hand, one may be interested in segregating from a large group of persons tested a smaller group which, on the average, will excel the larger group in whatever trait is being tested. This is, in fact, the situation that confronts the employment manager. He is willing to allow the tests to induce him to place upon a job a few individuals who will fail

on the job and to reject (or place upon some other job) a few who, if placed upon that job, would have succeeded, *if on the whole his percentage of successful placements is higher with the tests than it is without them.* In other words, the employment manager is not so much interested in every strike resulting in a home run as he is in improving his batting average. Under these circumstances the validity of the tests can be much lower. But one may still ask: How low can it be? A categorical answer to this question can be given, but the full significance of the answer will be clear only after a thorough study of the next section, which deals with the *selection ratio.* The answer is that a test will be valuable, no matter how low the coefficient of validity, if it indicates *some* relationship between test scores and the criterion; or, in statistical terms, if the coefficient of validity is at least four times its probable error. Often this rule will admit tests whose validity is as low as .30, or even lower. The use of tests with such low validity is sufficiently contradictory to much current thought among psychologists to warrant a fairly detailed justification for the conclusions reached above.

3. The selection ratio

Given a personnel test that has a validity coefficient indicating *some* relationship with the criterion, and given more employees to be placed than can be placed on the job in question, the functional value of the test to an employment manager depends upon the ratio of those placed to those tested who are available for placement. This has been referred to as the selection ratio.[5] An example will clarify the operation of this principle.

If a certain test is given to a large number of employees for whom a criterion of successfulness as employees is available, and if the scattergram of test scores against the criterion

[5] H. C. Taylor and J. T. Russell, "The Relationship of Validity Coefficients to the Practical Effectiveness of Tests in Selection: Discussion and Tables," *Journal of Applied Psychology*, XXIII (1939), pp. 565–578.

is plotted, the points will ordinarily fall into an oval-shaped area somewhat similar to the oval in Figure 15. The higher the coefficient of validity, the narrower will be the oval; and the lower the validity, the more nearly the oval will approach a circle. A validity coefficient of approximately .60 will result in a scattering of scores approximately covering the oval area shown in Figure 15. Now, if employees are placed

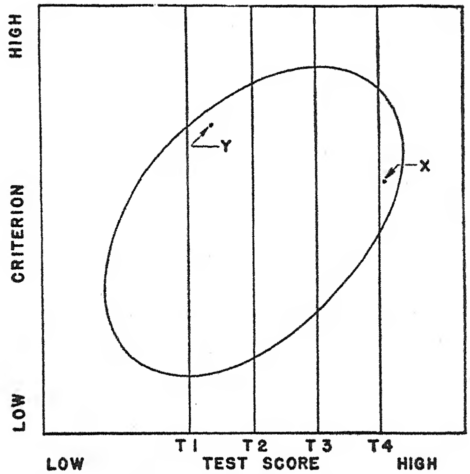


FIG. 15—Effect of shifting the critical score required of applicants on average criterion score of employees hired.

without regard to test scores, their criterion scores will be the average of all individuals falling within the oval. If only those are placed upon this job who have test scores as high as or higher than T_1, those not placed on the job will clearly have, on the average, lower criterion scores than the group as a whole, and those placed will accordingly be higher in their criterion scores, on the average, than the group as a whole. A still higher average criterion score for the group placed can be achieved by setting the critical test score at T_2. By moving the critical score to T_3, T_4, or even higher, still more favorable placements, according to average criterion score, can be made.

If a given number of persons, say 60, are to be placed, any one of the conditions mentioned above may exist; which one exists will depend upon the selection ratio that is utilized: that is, the ratio of the number placed to the number tested. Suppose we work with a ratio of 1.00: that is, all those tested are placed. In this case, the distribution of test scores will be over the whole range of possible test scores; the criterion scores will be over the whole range of possible criterion scores; and the test will contribute nothing whatever to the efficiency of the placement procedure. Now suppose that we test 80 individuals and place the 60 who score highest on the test, either not hiring the 20 who score lowest or placing them on some other job. We thus reduce those placed to 75 per cent of those tested, or reduce the selection ratio to .75. Under these conditions, we will place on this job only individuals who test at least as high as T_1, and the average criterion scores of those so placed will clearly be higher than the average of the group as a whole. By testing 120 persons and placing the 60 who score highest on the test, the selection ratio will be reduced to .50 and only individuals to the right of T_2 will be placed. The average criterion score of this group will not only be higher than that of the whole group, but will also be higher than that of the group placed when the critical test score was at T_1. Thus, by increasing the number tested before the 60 to be placed are identified, the selection ratio will be decreased with a continuous increase in the average criterion score for the group of 60 finally placed. If, for example, the plant is expanding so greatly that 600 new employees could be tested before 60 are selected for location upon this particular job (or if the labor market were such that 600 applicants were tested before 60 were hired for this job), the selection ratio would be decreased to .10, only those testing at least as high as T_4 would be placed upon the job, and the average criterion score of the group of 60 placed under these circumstances would be much higher than the score of 60 placed under any larger selection ratio.

The foregoing discussion is of course based on the assumption that the placement of employees is successful in proportion to the average success of the employees placed. Anyone can readily see that even working with a selection ratio of .1, some individuals (like X in Figure 15) will be placed who will be poorer according to the criterion than a few other individuals (like Y in Figure 15) who have not been allocated to this job. But if one is willing to measure the success of the testing program by average results rather than by individual cases, the results will be more and more favorable as the selection ratio is decreased.

We have already stated that psychologists dealing with vocational guidance and individual consultation are usually more interested in making accurate individual predictions than in making group predictions. Most of these psychologists have tended, therefore, to evaluate a test almost entirely in terms of its validity coefficient. They have stressed the fact (which is unquestionably true for individual prediction) that there is no substitute for high validity: that if two tests have been validated against the same criterion, and one has a higher validity coefficient than the other, there is no way to make the one having the lower validity serve as well as the other. The main point of the discussion is that in group testing, where one is interested in average rather than individual results, one can make the test with the lower validity perform as well as the other *by sufficiently reducing the selection ratio.*[6] In other words, in group testing, *a reduction of the selection ratio is a substitute for high validity.* This statement does not mean that this substitute will work if the test has no validity at all, but it does mean that if the test has any significant validity, however small, it is possible for the employer to get the same functional value from it that he could get from a test of any validity, however high, if he is able sufficiently to reduce the selection ratio.

[6] *Ibid.*, also Clark L. Hull, *Aptitude Testing* (World Book Company, 1928). See footnote on page 276.

A practical objection to the principle of increasing the efficiency of a test by decreasing the selection ratio may be raised, namely, that there is a limit to the number of applicants who can be tested before the desired number are placed. It is true that, for numerous reasons, an employment manager is seldom able to test 600 or even 200 men before 60 are placed. But it should be remembered that a ratio of 10:1 can be achieved by placing one person out of ten tested just as well as by placing 60 out of 600 tested. One does not need to wait for a great expansion in hiring before advantage can be taken of a reduced selection ratio. *A reduction in the selection ratio can be utilized whenever two or more employees are being placed on two or more different jobs, if tests of some validity are available for each of the jobs.*

The objection may also be raised that advantage cannot be taken of a reduced selection ratio unless there are more applicants than there are jobs to be filled. This is true if all persons employed are to be placed upon the same job. But almost always an expansion of plant personnel involves hiring for several jobs, not just for a single job. Therefore, the advantage of a reduced selection ratio usually can be achieved even when, as in a period of emergency production, there is difficulty in getting *enough* applicants to fill the jobs. Even when all applicants are hired, placements can be made on various jobs in such a way as to take advantage of individual differences by means of reduced selection ratios.

4. Percentage of present employees considered satisfactory

A further factor which affects the efficiency of a personnel test in a given employment situation is the percentage of present employees who are considered satisfactory. This factor may be made clear by reference to Figure 16. Suppose we are working with a test having a validity coefficient such that the employees tested fall into the oval-shaped area. Suppose, further, that we are working with a selection ratio of .5—that is, only persons falling in the oval-shaped area to

the right of T_2 will be placed on the job. If 50 per cent of the present employees are satisfactory, any increase over this amount in the percentage of satisfactory employees placed that can be achieved by using the test is a gain. Under these conditions, the ratio of satisfactory employees among those placed to the total of those placed would be the ratio of the number of individuals falling to the right of line T_2 and above C_2 to all persons falling to the right of line T_2. This ratio

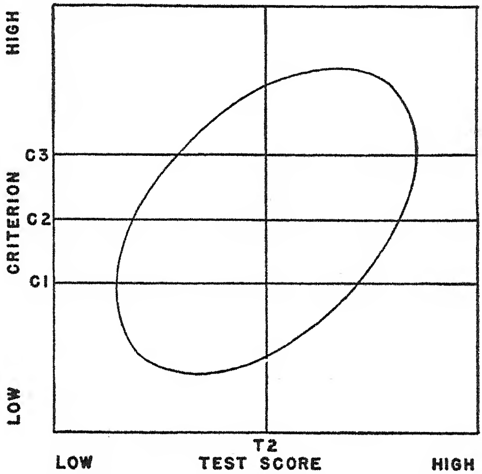


FIG. 16—Variation of efficiency of an employment test with differences in percentage of present employees who are considered satisfactory.

would clearly be higher than .50, and the amount by which it exceeds .50 would be indicative of the functional value of the test under the conditions discussed.

If all conditions named above remain the same except that previous employment methods have resulted in 75 per cent satisfactory employees, then the criterion separation line of the successful and unsuccessful employees would be C_1 and the percentage of satisfactory employees placed by means of the test would be the ratio of the individuals to the right of line T_2 and above C_1 to all persons to the right of line T_2. In the latter case, a larger percentage of employees hired will

be satisfactory than in the former case, even though the test, selection ratio, and other controlling factors remain the same. In other words, if everything else is equal, the smaller the percentage of present employees who have been placed satisfactorily without tests, the larger will be the percentage increase of satisfactory employees when employees are placed by means of test results. This may be illustrated by an

TABLE 8

INCREASES IN PERCENTAGE OF SATISFACTORY EMPLOYEES PLACED ON A JOB OVER VARIOUS ORIGINAL PERCENTAGES OF SATISFACTORY EMPLOYEES WHEN A TEST WITH VALIDITY COEFFICIENT OF .50 IS USED WITH A SELECTION RATIO OF .50

A Percentage of Satisfactory Employees Placed on the Job *without* the Test	*B* Percentage of Satisfactory Employees Placed on the Job *with* the Test	Difference in Percentage Between Columns *A* and *B*	Percentage of Increase of Values (*B*) over Values in (*A*)
5	9	4	80
10	17	7	70
20	31	11	55
30	44	14	47
40	56	16	40
50	67	17	34
60	76	16	27
70	84	14	20
80	91	11	14
90	97	7	8

example. Suppose we have available a test with a validity coefficient of .50 and are working with a selection ratio of .50. Table 8 shows the increase due to test in percentage of satisfactory employees which might exist prior to the use of the test. The values in Table 8 were obtained from the Taylor-Russell tables reproduced in Appendix B. If only 5 per cent of employees placed by traditional means are successful, then the expected increase to 9 per cent represents an 80 per cent

increase in the number of satisfactory employees placed by the test, under the specified conditions of test validity and selection ratio. For larger percentages of satisfactory employees that have been achieved without the test the percentage of increase achieved by using the test becomes increasingly smaller. If 90 per cent of employees placed by traditional means have been successful, the increase of this percentage to 97 per cent by the test, used under the specified conditions, results in an improvement of only 8 per cent in the number of employees satisfactorily placed.

The general conclusion is that, other things being equal, the more difficult it has been to find and place satisfactory employees without using test procedures, the greater the gain one may expect from a testing program.

The purpose of the foregoing discussion is to make clear the fact that several factors, each relatively independent of the others, operate to determine the functional value of a selection test to an employment manager. If the employment man knows these factors—that is, if he knows the validity of his tests, the selection ratio with which he is working, and the percentage of present employees considered satisfactory—he can predict definitely just how much he will improve the placement process by using the test. And if this amount of improvement is not satisfactory, he can improve his placement by almost any reasonable amount if the selection ratio can be decreased.

Figure 17 reproduces a chart that shows how the percentage of employees selected who will be successful is determined by the validity of the test and the selection ratio. This chart deals with an employment situation in which 50 per cent of present employees are considered satisfactory. The base line in this figure gives the selection ratio, and each of the curves plotted indicates a different test validity. It will be seen that by using a test with a validity of .90, and by reducing the selection ratio to .60, the percentage of satisfactory employees placed will be raised from 50 per cent to 77 per cent. It will

also be noted that a corresponding increase to 77 per cent in the number of satisfactory employees will be achieved by a test with a validity of only .50 if the selection ratio is decreased to .20.

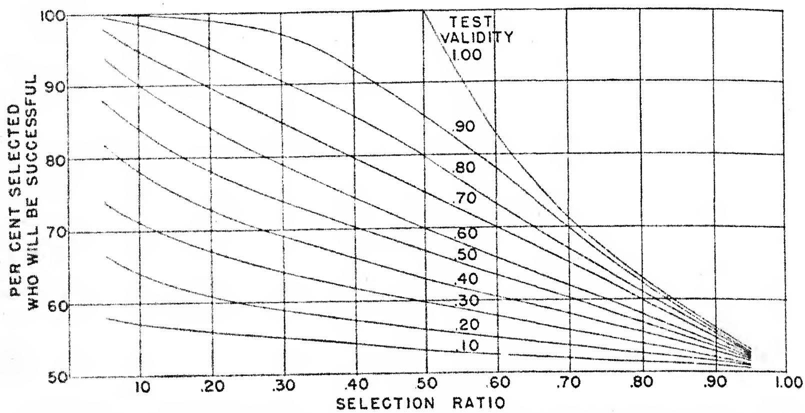


FIG. 17—Effect of test validity and the selection ratio upon the working efficiency of an employee selection test.

The Taylor-Russell tables,[7] reproduced in Appendix B, make it possible to determine what percentage of employees hired will be satisfactory under different combinations of test validity, selection ratio, and percentage of present employees considered satisfactory.

The use of these tables may be made clear by an example. Suppose an employment manager has a test with a validity coefficient of .40, has twice as many applicants or employees available as there are jobs to be filled, and is placing in a department where 30 per cent of the present employees are considered satisfactory. Looking in the lower half of the table on page 522 (entitled "proportion of present employees considered satisfactory = .30"), we find in the row representing a validity coefficient of .40 and in the column representing a selection ratio of .50, the value .41 where the indicated row and column cross. This means that under the conditions

[7] Taylor and Russell, *op. cit.*

specified 41 per cent of the employees placed will be satisfactory instead of the 30 per cent attained without the test. If conditions are such that the selection ratio may be still further reduced, the same test will place a still higher percentage of successful employees. For example, if only the highest 10 per cent of the persons tested are placed on the job, the percentage of satisfactory employees will be raised to 58 per cent, or nearly double the percentage of satisfactory employees placed without the test.

5. Combining tests into a battery

No single test will measure all of the capacities or abilities required on any job. Even the simplest of jobs is complex if one considers the combination of capacities or abilities required of a person who is to remain on the job and to do it well. Johnson[8] has pointed out that aptitude for any job consists of a syndrome of abilities and that one needs all of these to be successful. This fact makes it desirable, and in some cases necessary, to use a battery of tests rather than a single test. By means of statistical methods, the results of several tests can be combined into a composite score so that each is weighted to give the maximum correlation between the battery test score and the criterion. An example of the use of such combinations and weighting may be found in a set of tests worked out for placing menders in a hosiery mill. A number of tests were given to one hundred employees on this job. For each employee, data were obtained on age and experience as well as average hourly earnings for the twelve-week period preceding the administration of the tests. The correlations between several of the tests and the earnings criterion are given in Table 9.

It will be noted in Table 9 that the maximum correlation of any individual test with earnings was .27, but that a battery made up of all three correlated .35 with the same criterion.

[8] H. M. Johnson, "Some Neglected Principles of Aptitude Testing," *American Journal of Psychology*, XLVII (1935), pp. 159–165.

In obtaining the composite score the following formula was used: Composite Test Score = 12 (Hayes Pegboard) − 4 (Purdue Hand Precision Test) − 2 (Finger Dexterity Error

TABLE 9

Correlations between Several Dexterity Tests and Earnings of Menders in a Hosiery Mill

(Effect of age and experience statistically eliminated)

	Correlation with Earnings
Purdue Hand Precision Test*	.27
Finger Dexterity Test (Error Score)*	.18
Hayes Pegboard*	.16
Composite Score of Three Above Tests	.35

* Details of construction and administration of these tests are given in Chapter 5.

Score). The constants by which the raw test scores are multiplied in this formula result in the best combination of the tests to make a prediction of job success from a combina-

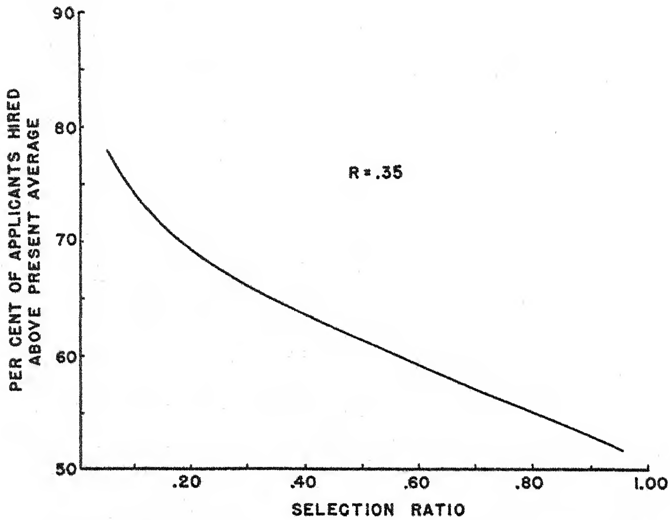


Fig. 18—Variation in per cent of employees who will be above the present average of employees when different selection ratios are used with a test having a validity coefficient of .35.

tion of test scores. The usefulness of a test battery that correlates to the extent of .35 with a criterion may be inferred from Figure 18, which shows the percentage of employees

placed by this battery who will be above the average of present employees when different selection ratios are used in hiring or placement. For example, if the selection ratio could be reduced to .10, approximately 70 per cent of the employees

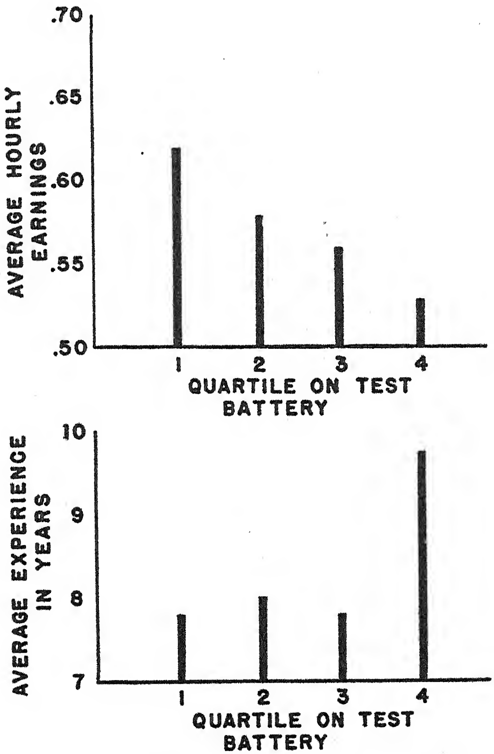


Fig. 19—Average earnings and average experience of four quartiles of "seamers" divided according to composite score on a battery of dexterity tests.

placed would be above the present average. The value of the testing procedure in this case is definitely enhanced by the use of several tests in combination, since no single test, of those tried, gave as high a validity coefficient as did the battery as a whole.

The way in which a battery of tests may be expected to

function in selecting the highest earners on a manipulative job is illustrated in Figure 19. To obtain this figure, another group of 150 hosiery workers (seamers) were first divided into four quartiles according to their composite score on two tests (Purdue Hand Precision and error score of the O'Connor Finger Dexterity) which had been found related to production on this job. For each quartile of employees, the average earnings and average experience were computed and plotted in Figure 19. It will be noted that there is a progressive decrease in average earnings from the best to the poorest quartile, in spite of the fact that the poorest quartile of employees on the test had a greater amount of experience, on the average, than any of the other quartile groups. Graphic representation of this type is possible, of course, only to the extent that no marked differences in experience exist among the groups compared. For example, if the low-testing group had two or three times as much average experience as the high-testing group (which actually occurred in the case of the menders previously mentioned), one could not make a direct comparison of the earnings of the different groups.

Further need for using a battery of tests, rather than an individual test, in many cases arises from the fact that since any single test will at best cover only one of several qualifications that are necessary on the job, a battery is required in order to obtain a more complete picture. The general appearance of the scattergram obtained by plotting test scores against a criterion often, in itself, indicates that a test, even a good test, covers only one phase of the job requirements. The scattergram often shows that persons testing high on the test may either be high or low on the criterion, whereas those testing low on the test are almost always low on the criterion.

An example of this tendency is shown in Table 10. The scattergram in this table shows the relation between the learning cost of 35 loopers in a hosiery mill and their scores on the Purdue Grooved Pegboard Dexterity Test (see page 130).

The minimum wage law now in effect requires that every employee be paid a certain minimum wage regardless of whether his piece-rate earnings reach this amount. A slow learner will thus have to be paid a large amount of "minimum make up" before she is earning her wage, whereas a fast learner will reach the legal minimum more quickly. Thus one indication, and a rather important one, of the desirability of an employee is the total amount of money that the company must contribute in minimum make-up wages while the employee is learning the job. It will be noted in Table 10 that this amount varies from $15.00, in the case of very rapid learners, to more than $60.00 in the case of slow learners.

TABLE 10

SCATTERGRAM SHOWING RELATION BETWEEN LEARNING COST TO COMPANY OF 35 LOOPERS IN A HOSIERY MILL AND SCORES AT TIME OF EMPLOYMENT OF THE PURDUE GROOVED PEGBOARD DEXTERITY TEST

LEARNING COST TO COMPANY	SCORE ON GROOVED PEGBOARD DEXTERITY TEST				
	60 seconds or over	55–59	50–54	45–49	40–44
$15–$24			2	2	5
$25–$34		1	3	3	
$35–$44			2		3
$45–$59	1			3	
$60 or over	2	3	3	1	1

The grooved pegboard is scored in time required to fill the board (best two trials out of four). The shorter the time, the better the score.

Table 10 makes several points clear. Employees with a favorable test score definitely tend to show a lower *average* learning cost to the company than do employees with a poor test score. In this case the correlation between test score (speed of filling the grooved pegboard) and economy in learning cost was .48. This figure may therefore be taken as the validity coefficient of the test for the job in question according to the learning-cost criterion.

But another characteristic of the scattergram shown in Table 10 that is commonly found in data of this sort is the *triangular* rather than the *oval*-shaped area into which the employees fall. Several employees who tested high actually turned out to be slow learners. However, no individuals with poor test scores turned out to be rapid learners. The significance of this finding is that the test in question apparently is measuring one essential requirement of a rapid learner, but only one requirement. An employee who lacks this requirement is practically certain to be a slow learner. But an employee who tests high, that is, who has an abundance of this requirement, may still be a slow learner if she lacks certain other basic requirements for the job. The triangular shape of the scattergram thus indicates the existence of a hierarchy or syndrome of capacities required for this job.

A question may be raised as to whether the reasoning on pages 66–75, in which the significance of the selection ratio concept was discussed, is valid when the scattergram is triangular rather than oval in shape. This reasoning applies equally well regardless of the shape of the scattergram as long as a vertical line drawn through the distribution at any point will divide the individuals plotted so as to give a higher average criterion score to those on the right of the line than to those on the left. This situation is certain to occur when a positive correlation between test scores and criterion exists. Thus, the rather common existence of a triangular scattergram does not invalidate the importance of the selection ratio; but it does quite definitely add to the evidence that a hierarchy of traits is basically necessary for satisfactory performance on most jobs.

6. Adequate training of testers

The impression that the tester makes upon applicants or employees is of vital importance to the success of a testing program. This fact, obvious to the experienced employment manager, has often been underemphasized or even completely

overlooked by psychologists themselves. "Everyone has shoes but the shoemaker's wife," says an old proverb. The professional psychologist often becomes so absorbed in the statistical phases of a test program that he forgets the importance of personal and industrial relations in actual test administration. Often the test administrator spends more time with an applicant than anyone else in the employment office. First impressions are lasting ones, and applicants who are employed may long remember the impressions of the company formed the first day at the plant. And the rejected applicants—upon whose reactions the industrial relations of the company within the community are also dependent—have no further opportunity to acquire a different opinion of the company. The man who administers the tests should be understanding, sympathetic, and courteous. He should give the applicant a feeling of importance. If the test results do not justify employment, the applicant should be made to understand that, although he is not properly adapted or trained for the job now open, he may be quite qualified for some other job at a later time or for some job now open in another plant. The applicant should be given an understanding of the fact that he himself would not profit from employment on a job on which he would be likely to fail.

To explain such matters to rejected applicants is neither easy nor routine. It must be done individually, thoroughly, and sincerely. Many of the qualities of a successful salesman are of real value to the test administrator in industry. Testing human beings is not the same as testing materials or processes. People *react*, either favorably or unfavorably, to a test situation. It is one job of the tester—indeed, one of his most important jobs—to be sure that they react favorably. Others may have prepared adequate tests; light, roomy, and attractive rooms may be available in which to give the tests; but unless the test administrator creates a favorable reaction to the testing program, the full advantage of this tool for employee selection and placement is not likely to be realized.

4

Mental Ability and Mechanical Aptitude Tests

For several reasons, we shall consider tests of mental ability before we consider tests in such areas as mechanical aptitude, dexterity, or trade ability. It is hoped, however, that such prior consideration of the field of mental ability will not give a wrong impression of the relative importance of mental ability tests. We do not consider intelligence tests first because they are more important than other tests, nor because most jobs demand persons of high intelligence. We consider this field first because probably no other area of psychological testing has been so thoroughly explored. As a result of this exploration we know the fields in which mental ability tests are of definite value and also the fields in which they offer little or no promise. Given a job description, therefore, we can predict with reasonable assurance whether a mental ability test will be of value in allocating employees to that job and what test or type of test is most likely to "come through" for the job. But it should be kept in mind throughout this discussion that for many jobs in modern business and industry intelligence or mental ability tests now available give no correlation whatever, either positive or negative, with job success. Later we shall discuss a number of jobs of this type and summarize certain studies that suggest other types of tests that have been found more satisfactory than mental ability tests in those particular areas.

Typical Mental Ability Tests

Otis Tests[1]

The Otis Self-Administering Tests of Mental Ability are typical of the more widely used and thoroughly standardized tests in this field. This series is called "tests" rather than "test" because it consists of four equivalent forms of a higher examination, designed for high-school students and college freshmen, and an additional four equivalent forms of an intermediate examination, designed for grades four to nine. These examinations are modeled after a group test of mental ability designed by Otis for use in a large commercial establishment. They are made up of logical and arithmetical problems, beginning at a very easy level, such as:

Which one of the five words below means the opposite of north?
1. pole, 2. equator, 3. south, 4. east, 5. west. ( )

The number of the correct answer is placed in the parentheses at the right. Among the special features which the series embodies are: the principle of self-administration, which eliminates the need for a trained examiner; a simplified scoring system; a variety of test material; separate norms for a twenty- or thirty-minute time limit; and a simplified chart for computing percentile scores from raw scores on the test.

A person's percentile score indicates that percentage of persons in the group on whom the test was standardized who are at or below the score of the persons tested. Thus if one does better on a certain test than 20 per cent of the standardizing group, he is said to have a percentile score of 20 or to be at the twentieth percentile. If he excels 90 per cent of the standardizing group he is said to have a percentile score of 90.[2]

---

[1] Arthur S. Otis, Otis Self-Administering Tests of Mental Ability (World Book Co., 1922).

[2] The computation of percentile scores is discussed in Appendix A, p. 505. A detailed treatment of the meaning and computation of percentiles is given by H. E. Garrett, *Statistics in Psychology and Education*, Third Edition (Longmans, Green and Company, 1947).

The Otis tests also provide for conversion of raw scores into equivalent I. Q. (intelligence quotient) values. The I. Q. is a measurement that is used widely in school testing of young children, and has become a somewhat common term in popular writing. However, since it is used very little in industrial testing programs (because it is not a particularly serviceable measurement when used with scores of adults, and because of a certain feeling of resentment that it arouses in many applicants and employees), we shall not explain its technical derivation in this discussion.[3]

The Otis tests are in common use in numerous industrial and business employment offices, and a number of investigations that have made use of these tests will be summarized. (See page 89.)

Wonderlic Personnel Test[4]

This test is an adaptation of the higher form of the Otis Self-Administering Test and is particularly adapted to the needs of business and industry. The word *personnel* rather than *mental ability* or *intelligence* is used in the title to avoid the negative reaction that many applicants and employees have to a test that obviously deals with mental ability. The adaptation consisted in selecting from the original Otis Test those items that were found to differentiate most markedly superior from inferior employees on various types of industrial and business jobs. The result of this careful selection of items is that for many industrial jobs the revision differentiates more satisfactorily than does the original form, and that the test may be given in twelve minutes instead of the twenty or thirty required for the original test. The general form of

---

[3] The interested student will find a description of the meaning and derivation of the I. Q. in any elementary textbook of psychology, such as J. Tiffin, F. B. Knight, and E. J. Asher, *The Psychology of Normal People* (D. C. Heath and Company, 1946).

[4] E. F. Wonderlic, Personnel Test, Form D (E. F. Wonderlic, 919 N. Michigan Ave., Chicago).

the questions on the Wonderlic test is similar to that of the Otis test. The test is begun with such items as:

> Of the five things below, four are alike in a certain way. Which is the one not like these four?
> 1. smuggle, 2. steal, 3. bribe, 4. cheat, 5. sell. ( )

The correct answer of course is *sell*, the other four choices all being alike in the fact that each involves an element of dishonesty.

The Adaptability Test

The Adaptability Test[5] measures mental adaptability or mental alertness. The test has a fifteen-minute time limit and is useful in helping to identify persons who should be placed on jobs that require rapid learning and/or the development of independent judgment. It also aids in identifying persons who do not readily adapt to new situations but who would be satisfactory (and often superior) employees on simple, routine jobs such as packing, inspecting, assembling, or in the operation of simple, repetitive machines.

The test was constructed specifically for industrial use. The questions were phrased in practical rather than academic terminology.

The introductory paragraph on the title page of the test was developed through consultation with industrial personnel men and actual try-out with applicants and employees. This introduction, which is reproduced below, makes no mention of mental ability, intelligence, or the I. Q., and represents an effort to put the individual taking the test at ease:

> Some jobs require figuring—such as adding, subtracting, multiplying, and dividing—while others require writing reports or answering letters, and still others can be done well by people who are not particularly apt with figures or words. This test will help in determining how well you can handle jobs that require these abilities.

---

[5] This test is distributed by Science Research Associates, 228 S. Wabash Ave., Chicago, Ill.

Do as well as you can on this test, but do not worry about it. Remember that you may be well qualified for certain jobs that require training or skills different from those covered in this test.

(*Reproduced with permission of Science Research Associates*)

The Revised Beta Examination[6]

This is a group test of mental ability that is strictly non-verbal; that is, it does not call for ability to read. It is particularly useful in measuring persons who have a foreign language background or whose previous work or training has been such as to penalize them on a test that involves reading and language.

The Psychological Examination of the American Council on Education[7]

This is another group test of mental ability which, because of its longer time limit of fifty minutes, has not been used so extensively in business and industry as has either of the foregoing tests. For some purposes, however, particularly college placement and vocational guidance, it has proved to be an excellent instrument. Several studies[8] indicate that it has also definite value for business and industry.

The preceding tests have been mentioned only as illustrative of the kinds of mental ability tests that are available. It is not our purpose here to attempt even a partial listing, to say nothing of an evaluation, of the many excellent measuring devices available in this field. *The Mental Measurements*

---

[6] C. E. Kellogg and N. W. Morton, Revised Beta Examination (The Psychological Corporation, 522 Fifth Ave., New York, N. Y.).

[7] L. L. Thurstone and Thelma Gwinn Thurstone, Psychological Examination of the American Council on Education (Science Research Associates, 228 S. Wabash Ave., Chicago, Ill.).

[8] S. R. Laycock and N. B. Hutcheon, "A Preliminary Investigation into the Problem of Measuring Engineering Aptitude," *Journal of Educational Psychology*, XXX (1939), pp. 280–289.

J. B. Rhinehart, "An Attempt to Predict the Success of Student Nurses by the Use of a Battery of Tests," *Journal of Applied Psychology*, XVII (1933), pp. 277–293.

Yearbook[9] gives a fairly complete listing of intelligence tests. For most of the tests listed, the *Yearbook* gives the title, a description of the group for which the test was constructed, the date of copyright or publication, whether the test is individual or group, the number of forms, the cost, the time required for administration, the author and the publisher, references to studies dealing with validity and reliability, and a brief evaluation of the test by one or more competent authorities. The industrial personnel man looking for a test that will serve some particular purpose can usually decide from the information given which test will be most likely to serve his purpose.

Another list of mental ability tests that have been used successfully in business and industry is reproduced in Appendix C on page 526.

What tests to use

Once an employment manager has decided to use intelligence tests, it is not uncommon for him to ask a psychologist, "What intelligence test do you recommend?" It is difficult, if not impossible, to give a categorical answer to this question. The test to use is the test that works, and all tests do not work equally well in all situations. Asking a psychologist what intelligence test he recommends is not unlike asking a physician what drug he recommends. The physician's recommendation depends upon a number of factors, which include what is wrong with the patient, what the drug is intended to accomplish, and the present physiological condition of the patient. The psychologist cannot answer the question "What test do you recommend?" by mentioning any specific test. The question can only be answered by summarizing the results of studies that have been obtained by using various tests and types of tests in industrial and business

---

[9] Oscar K. Buros, *The 1940 Mental Measurements Yearbook* (Rutgers University Press, 1940).

situations. The test to be used in any new situation should then depend upon the extent to which the new situation is parallel or equivalent to those situations that have been studied.

Mental Ability Tests in Clerical Selection

Although a number of studies indicate that it is often advisable to use a standardized intelligence test in the selection of persons for clerical jobs, it should be kept in mind that the term clerical is very broad. Under this category might be included regular clerks, comptometer operators, stenographers, secretaries, and persons on other related jobs. Even in any one subcategory—as stenographer—the importance of mental ability might vary all the way from being an essential requisite on the job to being of no importance at all. For example, a secretary who is expected to answer mail during the employer's absence should have a high degree of what we ordinarily refer to as tact, judgment, and discretion—the things that constitute mental ability. On the other hand, a secretary whose main function is to greet customers or guests in the outer office and to keep them happy and contented until the employer has time to see them does not need a high level of mental ability so much as she needs personal attractiveness, social intelligence, and the ability to carry on a light conversation. A standard mental ability test might be a very desirable instrument to use in selecting a secretary for the former type of work, though there would be little to recommend such a test for the selection of the receptionist. Each test, therefore, must be evaluated in terms of whether it is measuring something that the employee must possess if he is to do the job well. If this point is kept in mind, we can readily understand why some investigators have found appreciable correlations between intelligence and clerical ability while others have reported little or no correlation. The choice of a test to be used, then, should depend upon the type of clerical ability under consideration.

Typical of studies in this field is one made by McMurry[10] who found correlations varying from .34 to .57 between Otis Intelligence Test scores and job efficiency among bank employees. The employees considered were for the most part engaged at least partially in a type of clerical work that involved careful and accurate computation and machine operation. Pond and Bills[11] have reported a study which further shows the importance of intelligence tests in selecting individuals for certain types of clerical jobs. Their investigation shows that on the easiest jobs the greatest turnover is among the employees with highest test scores, whereas on the more difficult jobs the greatest turnover is among the employees testing lowest in mental ability. In other words, the more difficult clerical jobs call for employees with the type of ability that a mental ability test measures. If all the employees possessing such ability are placed upon such jobs, whereas those not possessing such ability are placed on the easier jobs, the total turnover of the employees on both types of jobs will be definitely reduced.

Shellow[12] reports some interesting correlations between intelligence test results and the ratings of ability of a group of stenographers. The employees studied were given two tests —an intelligence test and a test measuring proficiency in stenographic skills. The intelligence test contained proverbs, items of the completion type, the analogy type, and so on— all items of the type included in most standardized tests of intelligence. It may safely be assumed that the particular test used measured much the same aspect of mental ability that is measured by a standard intelligence test. The stenographic test was an ordinary test of proficiency in

---

[10] R. N. McMurry, "Efficiency, Work-Satisfaction and Neurotic Tendency. A Study of Bank Employees," *Personnel Journal*, XI (1932), pp. 201–210.

[11] M. Pond and M. A. Bills, "Intelligence and Clerical Jobs. Two Studies of Relation of Test Score to Job Held," *Personnel Journal*, XII (1933), pp. 41–43.

[12] S. M. Shellow, "An Intelligence Test for Stenographers," *Journal of Personnel Research*, V (1926), pp. 306–308.

stenography of the type discussed on page 153. The correlations reported by Shellow are as follows:

Intelligence Test versus Ranking	.73
Proficiency Test versus Ranking	.48
Rough combination of above tests	.59
Properly weighted combination of above tests	.83
Intelligence Test versus Proficiency Test	.12

These correlations show that, for the group of stenographers studied, intelligence as measured was even more important than stenographic proficiency in determining rated ability on the job. The results also show, as might be expected, that the combination of intelligence and proficiency in stenography makes for a higher ability on the job than either of the two taken alone. It is interesting to note that the rough combination of the intelligence and the proficiency tests gave a correlation with ranking of only .59, which was lower than the .73 between intelligence alone and the ranking. This result simply means that when two tests are to be combined for prediction the tests cannot be optimally weighted except in the light of the multiple correlations involved. As the table shows, when the optimal weighting had been obtained, the correlation was .83 between the combined score on the two tests and the ranking on the job.

In the selection of cashiers, Clarke[13] has reported correlations as high as .57 between productive efficiency as measured by transactions handled per day and scores on the Otis Test, a Sales Checking Test, a Change Making Test (speed and accuracy), manual dexterity, and the Bernreuter Scale. The Bernreuter Scale, which will be discussed later (see page 166), could hardly account for all of the relationship found. It seems clear, therefore, that the abilities tapped by the mental ability test are of primary importance in determining the job efficiency of the cashiers studied. A study by Stevens and

[13] W. V. Clarke, "The Evaluation of Employment Tests," *Personnel*, XIII (1937), pp. 133–136.

Wonderlic[14] reports the relationship between scores on the Otis Mental Test and ability to handle office detail. Their study indicates that the employees who missed most items on the Otis Scale are the ones most criticized by company supervisors.

Tiffin and Lawshe[15] found correlations ranging from .40 to .65 between rated success of paper mill clerical employees and scores on the Adaptability Test. In one department where only 57 per cent of current employees were considered satisfactory, it was found that among those employees making scores of 25 or above on the Adaptability Test, 86 per cent were satisfactory employees.

Hay[16] reports a correlation of .56 between Otis Test scores and production of machine bookkeepers. In another study, Hay[17] reports a correlation of .35 between rated ability of clerical employees and Otis Test scores.

It would seem safe to conclude, on the basis of numerous studies such as those here summarized, that an employment manager can safely use some sort of mental ability test in selecting employees for most types of clerical jobs.

Mental Ability Tests in Selecting for Other Jobs

A number of other investigations suggest the usefulness of mental ability tests for selecting individuals for various types of jobs. Harrell[18] reports a correlation of .37 between Otis scores and the ability of cotton mill supervisors. Copeland[19]

---

[14] S. N. Stevens and E. F. Wonderlic, "The Relationship of the Number of Questions Missed on the Otis Tests and Ability to Handle Office Detail," *Journal of Applied Psychology*, XVIII (1934), pp. 364–368.

[15] Joseph Tiffin and C. H. Lawshe, Jr., "The Adaptability Test: A Fifteen Minute Mental Alertness Test for Use in Personnel Allocation," *Journal of Applied Psychology*, XXVII (1943), pp. 152–163.

[16] Edward N. Hay, "Predicting Success in Machine Bookkeeping," *Journal of Applied Psychology*, XXVII (1943), pp. 483–493.

[17] Edward N. Hay, "Tests in Industry. I. Their Proper Use," *Personnel Journal*, XX (1941), pp. 3–9.

[18] W. Harrell, "Testing Cotton Mill Supervisors," *Journal of Applied Psychology*, XXIV (1940), pp. 31–35.

[19] H. A. Copeland, "Validating Two Tests for Census Enumeration," *Journal of Applied Psychology*, XXI (1937), pp. 230–232.

reports a low but positive correlation between a test battery made up of the Otis Test and the Minnesota Clerical Test and the ratings of supervisors, clerks, and enumerators in the census department. Wadsworth[20] reports that among the entire force of a utility plant, intelligence correlated .68 with "man-to-man" estimates by supervisors. He also reports that after testing was adopted as a part of the employment procedure, the percentage of unsatisfactory employees hired dropped from 29 per cent to 5.5 per cent. Sheddan and Witmer[21] report that a battery of tests consisting of the Ohio State Psychological Test (an intelligence test), the Moss Social Intelligence Test, the Thurstone Personality Test, and a Relief Attitude Scale, was found to have a correlation of .72 with job efficiency of 61 relief visitors, as measured by their scores on a test of technical information about their duties and by merit rating of their performance by superiors. Holliday[22] reports that an investigation of the selection of apprentices for the engineering industry revealed that standard intelligence tests definitely improved selection. Monro and Raphael[23] found that a ten-minute test designed to measure intelligence, arithmetical accuracy, attention to instructions, and tact significantly differentiated "good" from "fairly good" salesgirls. Laycock and Hutcheon[24] found that a test battery consisting of high-school grades, the American Council Psychological Examination, Cox Mechanical Aptitude Test (see page 109), Minnesota Paper Form Board (see page 105), and interest in physical science gave a multiple correlation of .66 with success in engineering.

---

[20] G. W. Wadsworth, "Tests Prove Worth to a Utility," *Personnel Journal*, XIV (1935), pp. 183–187.

[21] B. R. Sheddan and L. R. Witmer, "Employment Tests for Relief Visitors," *Journal of Applied Psychology*, XXIII (1939), pp. 270–279.

[22] F. Holliday, "An Investigation into the Selection of Apprentices for the Engineering Industry," *Occupational Psychology*, XIV (London, 1940), pp. 69–81.

[23] M. S. Monro and W. S. Raphael, "The Value of a Short Test for the Selection of Workers," *Human Factor*, VI (1932), pp. 244–246.

[24] Laycock and Hutcheon, *op. cit.*

Another area in which the use of a mental ability test has been found very helpful is in the selection of teletype operator trainees. It is customary for the operating company to send trainees for this job to a special school where a carefully prepared program of instruction is given. The trainees are graded periodically by means of standardized and objective tests which cover both job performance and related information. Trainees are certified as operators only after successfully passing this course.

In a recent investigation, 219 teletype trainees were given the Adaptability Test, previously described, shortly after being selected for training. The follow-up method of test validation was employed to determine the relation between test scores and success in the training course. The results are summarized in Table 11.

TABLE 11
RELATION BETWEEN SCORES ON THE ADAPTABILITY TEST AND SUCCESS IN A TRAINING COURSE FOR TELETYPE OPERATORS

Score on Adaptability Test	Number Passing Course	Number Failing Course	Total Number	Per Cent Passing Course
22 or above	27	1	28	96%
16–21	66	4	70	94%
10–15	83	10	93	89%
9 or below	17	11	28	61%

The cost to the company for each trainee (which includes transportation to and from the school, board and room, salary during that training, and the pro-rata cost per trainee of operating the school) amounts to approximately \$500 per trainee. The investigation cited above shows that from every 100 trainees who score 22 or above on the Adaptability Test, the company will be able to certify 96 operators, the cost thus being $\frac{100 \times \$500}{96}$ or \$520.83 per operator. Among trainees scoring 9 or below on the Adaptability Test, of every

100 trained only 61 will be certified as operators, the cost thus being $\frac{100 \times \$500}{61}$ or \$819.67 per operator. The comparative costs show that *each of the operators in the low-score bracket costs the company \$298.84, or approximately 57 per cent, more in training expenses than do operators in the high-score bracket.*

Bolanovich[25] reports a study which further shows the value of a mental alertness test in selecting trainees for a course of study that requires study and application. His investigation deals with a series of tests given to a group of girl trainees referred to as the R.C.A. Cadettes—girls sent by the Radio Corporation of America to a specialized training course covering certain phases of traditional engineering subjects. The correlation obtained between Wonderlic Personnel Test scores and grades in the course was .50.

Mental ability tests also have been found highly advantageous as one tool to be used in the selection of supervisors. A rubber company recently promoted 70 operators to supervisory jobs. To train these men for their new duties and responsibilities, the company conducted a series of supervisory training classes, at an expense to the company of \$225 per man for the time spent in attending the classes. This amount did not include instructional costs or other expenses. During the first meeting of the supervisory training class, the men were given the Adaptability Test discussed above. The test scores were filed and a period of six months was allowed to pass while various criteria of success on the job of the new supervisors were investigated. At the end of the six-month period, it became apparent that the most satisfactory criterion of success available in this instance was whether or not the man was a good enough supervisor to be still on his new job six months later. Approximately one-fourth of the men were not, having quit or been demoted,

[25] D. J. Bolanovich, "Selection of Female Engineering Trainees," *Journal of Educational Psychology*, XXXV (1944), pp. 545–553.

transferred, or dismissed during the six-month period. In the case of the men no longer on the job, the cost of the supervisory training, as well as the other considerable expenses incurred in breaking in the new men, was borne by the company without return. Figure 20 shows the percentage of men

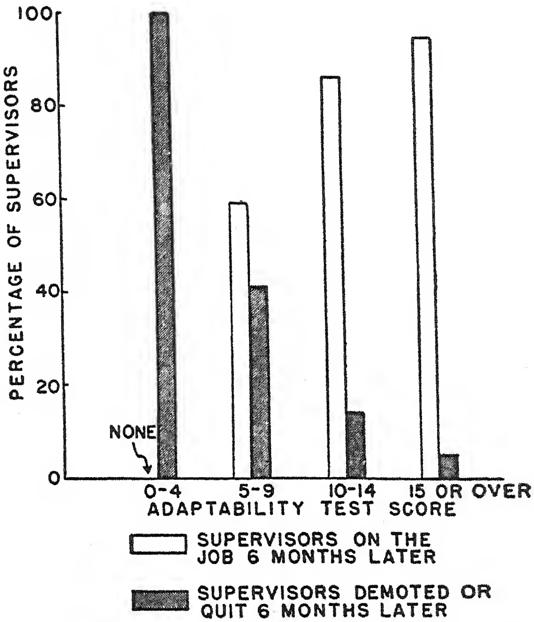


FIG. 20—Percentages of men promoted to supervisory jobs in a rubber company who were still on the supervisory job six months later.

still on the supervisory job six months later among those scoring at various levels on the Adaptability Test. These findings are in agreement with those reported by Moore,[26] who emphasizes the importance of using some form of mental ability test, in connection with other requirements, in the selection of supervisors.

---

[26] H. Moore, "Supervision. I. Selection," *Personnel Journal*, XX (1942), pp. 353–356.

Jobs in Which Success Is Not Related to Mental Ability

The preceding discussion of mental ability tests has involved a deliberate selection of investigations in which the tests have "come through." It should be pointed out and emphasized that in a good many investigations the mental ability tests did not come through. One example of a failure of a mental ability test to contribute anything of value with regard to predicting probable success of employees is an extensive study by the follow-up method of 749 employees hired for bench work in the Hawthorne Plant of the Western Electric Company.[27] After these employees had been on the job long enough to give a variety of indications of their success or failure on the job, including such things as promotion, average wage increase since hiring, merit ratings by supervisors, average tenure with the company, and freedom from accidents, the correlations were computed between the several criteria of successful job performance and the original test scores obtained at the time of hiring. Of four tests given at the time of employment, one was the Otis higher form of the mental ability test. The Otis scores showed practically no relation with later success or failure of the employees on their shop assignments, though it is of interest to note that this test indicated ability to progress to higher-level shop occupations or office jobs. In other words, the intelligence test was definitely an indication of "promotability." However, two tests dealing with dexterity, also given at the time of hiring, did show a definite relationship with the later success of the employees. These tests will be discussed on page 139.

Still other investigations have shown that mental ability tests sometimes show a negative relationship with success on the job. For example, Tiffin and Greenly,[28] in studying

[27] *Analysis of 1935–37 Experience in Selecting New Men for Shop Occupations.* Privately Printed Monograph (Western Electric Company, Hawthorne Plant, 1939).

[28] Joseph Tiffin and R. J. Greenly, "Employee Selection Tests for Electrical Fixture Assemblers and Radio Assemblers," *Journal of Applied Psychology*, XXIII (1939), pp. 240–263.

routine manipulative assembly, found a negative correlation between the Otis scores and both the actual amount of production and supervisors' ratings of productivity. Tiffin and Lawshe[29] found that the best inspectors of simple material made lower average Adaptability Test scores than did the poorest inspectors on the same job. Abel[30] reports a study of 84 girls from the Manhattan High School for the Women's Garment Trades who left school at seventeen years. The average Otis intelligence test score of the girls indicated feeble-mindedness or near feeble-mindedness, and yet after a period of 3.5 years, 55 per cent were steadily employed. Abel points out that among the factors contributing to their success were stable homes, ambition and self respect, and encouragement and patient treatment during the initial work period.

These studies are in accord with studies already referred to by Bills[31] and by Pond and Bills[32] showing that individuals who test high on mental ability usually should not be placed on routine, easy jobs, and that frequently persons scoring very low on mental ability tests consistently do better on simple, repetitive jobs than do persons whose mental ability is average or above.

General Recommendations Concerning Mental Ability Tests

After studying the preceding and similar investigations, one feels a strong temptation to generalize concerning the role of mental ability tests in business and in industrial employment offices. Such a generalization might be that mental

---

[29] Tiffin and Lawshe, *op. cit.*

[30] T. M. Abel, "A Study of a Group of Subnormal Girls Successfully Adjusted in Industry and the Community," *American Journal of Mental Deficiency*, XLV (1940), pp. 66–72.

[31] M. A. Bills, "Relation of Mental Alertness Test Scores to Positions and Permanency in Company," *Journal of Applied Psychology*, VII (1923), pp. 154–156.

[32] Millicent Pond and M. A. Bills, "Intelligence and Clerical Jobs. Two Studies of Relation of Test Score to Job Held," *Personnel Journal*, XII (1933), pp. 41–56.

ability tests are of value in selection and placement to the extent that the employees being hired are to be placed on jobs calling for adaptability, accuracy, carefulness, and the exercise of judgment in situations that are not likely to be exactly the same from day to day. At the other extreme, individuals testing average or below on mental ability tests would seem to be most suitable for jobs involving a completely repetitive or mechanical type of manipulation of a not-too-

	Decile	Finger Dexterity Minutes to fill board		Hand Precision Error score		Otis Test Raw score	
Highest 10 percent	1	5.85 or less		0-56	✓	58 or above	
	2	5.86-6.42		57-79		53-57	
	3	6.43-6.82		80-104		49-52	
	4	6.83-7.17	✓	105-144		45-48	
Average	5	7.18-7.49		145-170		43-44	
	6	7.50-7.81		171-232		39-42	
	7	7.82-8.16		233-273		37-38	
	8	8.17-8.56		274-279		33-36	
	9	8.57-9.13		380-719		27-32	
Lowest 10 percent	10	9.14 or longer		720 or over		26 or below	✓

FIG. 21—Psychograph of an operator 32 per cent above average production on manipulative work.

difficult nature. A question that will naturally be raised by the last statement is, "How low can an individual be in mental ability and still be able to do satisfactory work on a routine manipulative job?" It would seem that for many simple and routine jobs, individuals testing at the very bottom on standardized tests of mental ability are as well or even better adapted than persons testing at average or above. Figure 21 shows the test profile of an individual who, for the

two-month period preceding the administration of the test, had consistently been the highest-producing operator among forty-two operators on this job. The job was one that involved routine manipulative assembly. It will be noticed that while this employee is above average in dexterity, she is far below average in her Otis Test score. Indeed, one would

	Decile	Finger Dexterity Minutes to fill board		Hand Precision Error score		Otis Test Raw score	
Highest 10 percent	1	5.85 or less		0-56		58 or above	
	2	5.86-6.42		57-79		53-57	
	3	6.43-6.82		80-104	✓	49-52	
	4	6.83-7.17		105-144		45-48	
Average	5	7.18-7.49	✓	145-170		43-44	
	6	7.50-7.81		171-232		39-42	
	7	7.82-8.16		233-273		37-38	
	8	8.17-8.56		274-279		33-36	
	9	8.57-9.13		380-719		27-32	✓
Lowest 10 percent	10	9.14 or longer		720 or over		26 or below	

FIG. 22—Psychograph of an operator 27 per cent above average production on manipulative work.

conclude from a literal interpretation of this test score, in the light of the published norms for this test, that this employee is feeble-minded. However, it should be kept in mind that these norms were derived from the point of view of an individual's ability to succeed in adjusting himself to new and novel situations. The individuals who are lowest on this test might be most likely to succeed in jobs that do not call for such adaptation to new and novel situations. Figure 22 shows a psychological profile for another individual who was

very low on the Otis Test and yet who was close to the top in productivity on the job in question.

Conclusions cannot, of course, be reached from single cases such as these. Yet when one considers that for the group as a whole negative correlations were found between mental ability test scores and productivity, and that a few individuals who tested at the very bottom on the Otis Test consistently led the group in productivity on the job, one is forced to conclude that it may be a serious mistake to place persons scoring high in mental ability tests on jobs of this type, unless they are being given job experience in preparation for supervisory positions.

The management of many industries has often made the mistake, during times of depression when applicants are available in large numbers, of hiring only individuals who offer a certain minimum of education—for example, high-school graduates. When an employment manager hires only high-school graduates he is, in effect, hiring only individuals who are above a certain level of mental ability. The content of courses taught in high school has no relationship to an industrial plant assembly job. But we know that only individuals who on the average possess a certain minimum level of mental ability are able to pass these courses and be graduated from high school. Thus the selection of high-school graduates is equivalent to selecting only those individuals who make a certain minimum score on a mental ability test. In one department of a large industry which employed approximately 400 operators for a routine job, the policy for several years during the depression of 1930 to 1935 was to employ only high-school graduates. This hiring policy resulted in an unusually high turnover, since individuals with high mental ability left the job as quickly as they could find something better. The policy was accompanied by, and may actually have caused, a condition of unrest and dissatisfaction among the employees. It has been said that nearly every individual finds some sort of outlet for whatever ability he may possess.

It is not unreasonable to suppose that a large part of the unrest and dissatisfaction arising among the 400 employees in this department was due in no small measure to the fact that the high-school graduation requirement resulted in placing individuals on this job who in terms of ability were "above the job."

It should be re-emphasized at this point that the primary function of mental ability tests in industry is to help the employment manager in his placement of employees. Adequate placement means that some jobs should be filled by high-scoring individuals and that other jobs should be filled by low-scoring individuals. A mental ability test is simply another tool that the employment manager may use in placing every individual on a job that is equal to and not above or below his natural level of accomplishment.

Typical Group Mechanical Ability Tests

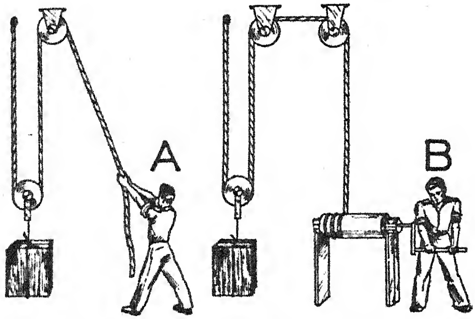
The purpose of mechanical aptitude tests

Mechanical aptitude tests have been developed to aid in selecting employees for jobs that require a mechanical "knack," such as the maintenance of machinery found in almost every plant, the operation and maintenance of conversion machinery in a paper mill, or the "setting-up" of production machinery before it is turned over to an operator.

Bennett Test of Mechanical Comprehension[33]

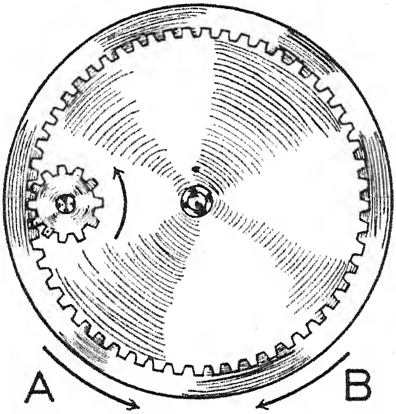
One of the best known tests in this field is the Bennett Test of Mechanical Comprehension, which was developed to accomplish the purposes mentioned above. The principle of the test is illustrated in Figure 23. The person tested is required to make a judgment from an examination of a drawing or schematic diagram. No mathematical or arithmetical computations are required, and the verbal or reading element

[33] G. K. Bennett, Test of Mechanical Comprehension, Form AA (The Psychological Corporation, 522 Fifth Avenue, New York, N. Y., 1940).



Which man can lift more weight?

Fig. 23—Two items from Form AA of the Bennett Test of Mechanical Comprehension.



If the small wheel goes in the direction shown, in which direction will the large wheel go?

is reduced to a minimum. As modern industry hires vast numbers of persons with practical mechanical experience but without formal education, the test fills a definite need in enabling an employment manager to measure mechanical ability among such persons. An illustration of the use of this test in a practical situation is given on page 117.

Detroit Mechanical Aptitudes Examination[34]

This test consists of several parts, some of which deal with the understanding of certain mechanical principles on a

[34] Detroit Mechanical Aptitudes Examination for Boys (Public School Publishing Co., Bloomington, Ill., 1928).

non-verbal level—such as the belt-pulley arrangement shown in Figure 24—and others with familiarity with tools and power machinery used in a wood shop or a machine shop. Figure 25 illustrates one item from a part of the test that deals with power machinery. Perhaps the major differences between the Bennett Test of Mechanical Comprehension and the Detroit Mechanical Aptitudes Test is that the latter is based to a much greater extent upon familiarity with actual tools and machines. This statement is not intended as a criticism

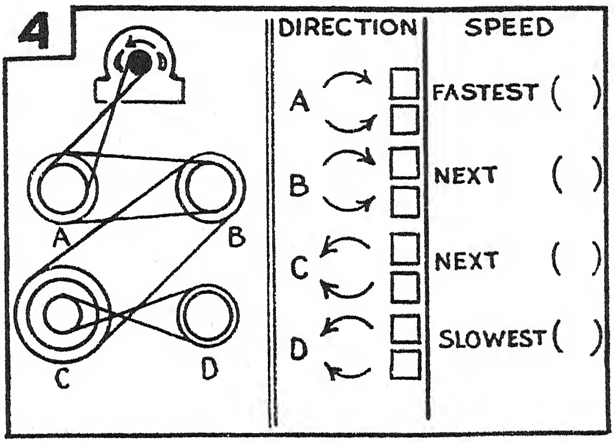


FIG. 24—An item from Part I of the Detroit Mechanical Aptitudes Test for Boys. The subject checks in the indicated blanks to show the direction and speed of the several pulleys.

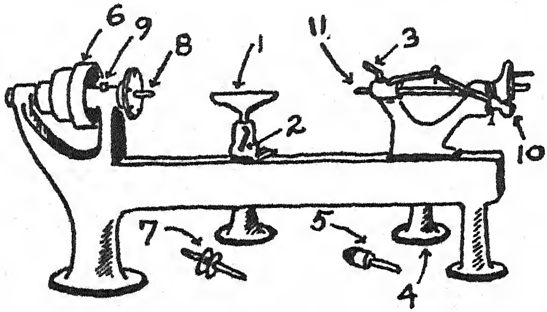
of either test. It means simply that the tests are likely to find their greatest degree of usefulness with different groups of applicants.

Stenquist Mechanical Aptitude Test[35]

This test is one of the oldest and perhaps one of the best known of all tests in this general field. It was first published in 1921 and at that time represented a real achievement in the application of group testing methods to the measurement of mechanical ability. An item from this test is illustrated in

[35] J. L. Stenquist, Mechanical Aptitude Test (World Book Company, 1921).

Figure 26. The Stenquist test, like the Detroit test, is based to a large extent upon familiarity with actual shop machinery. This characteristic is illustrated in Figure 26, which shows an overhead pulley assembly.



CARRIES BELT----------- (   )

ADJUSTS HEIGHT OF REST--- (   )

SUPPORTS TOOL--------- (   )

OILS BEARING---------- (   )

FASTENS TAIL STOCK CENTER- (   )

REVOLVES WORK--------- (   )

ADJUSTS TAIL STOCK CENTER- (   )

HOLDS WORK----------- (   )

FASTENS TAIL STOCK------ (   )

HOLDS DRILL----------- (   )

HOLDS EMERY WHEEL---- (   )

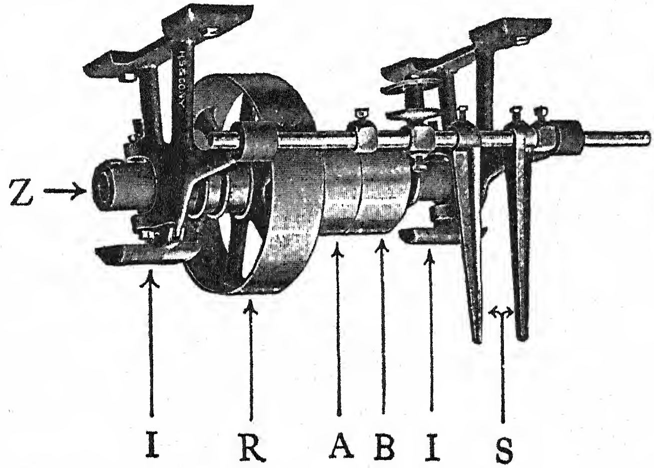
FIG. 25—An item from another part of the Detroit Mechanical Aptitudes Examination for Boys. The parentheses are to be filled in with the correct numbers.

O'Rourke Mechanical Aptitude Test[36]

This test, like the Stenquist test described above, measures knowledge of and familiarity with tools and operations involved in shop work. It is, therefore, basically an *achievement* test rather than an *aptitude* test. However, experience has shown that applicants who have the aptitude for mechanical achieve-

[36] L. J. O'Rourke, O'Rourke Mechanical Aptitude Test (The Psychological Institute, 1937).

ment ordinarily create opportunities for familiarizing themselves with shop equipment, and that such applicants are able, because of this experiential background, to obtain favorable scores on this and similar tests. This fact justifies the use of this type of test as an aptitude test for selecting applicants who are to receive additional training, such as apprentices.



1 Oil is kept from dripping on the floor by.......................... 1 ______
2 The belt is pushed back and forth by part......................... 2 ______
3 If pulley R is loose on the shaft, write L. If it is fastened to the shaft and turns with it, write F.............................. 3 ______
4 Does the belt which drives this machine run on pulley marked Z? Write Y is yes; N if no...................................... 4 ______
5 Look at the two pulleys A and B. If both turn with the shaft S, write B. If neither does, write N. If only one does, write O... 5 ______

FIG. 26—An item from the Stenquist Mechanical Aptitude Test. The questions are to be answered by filling in the blanks at the right.

An illustrative item from the O'Rourke Mechanical Aptitude Test is shown in Figure 27 on page 106.

Revised Minnesota Paper Form Board Test[37]

This is a unique test that deals with ability to visualize and assemble a set of blocks mentally. It largely eliminates

[37] Minnesota Paper Form Board Test (Science Research Associates, 228 S. Wabash Ave., Chicago).

SAMPLES

Each of the three pictures marked with a number is **used with** a picture at the right marked with a letter. Look at the picture marked 1. Then look at the pictures marked A, B, and C and decide which is **used with** 1. Write the letter of the picture which goes with 1, on the line marked 1 at the right of the pictures. Then find the picture that is **used with** picture 2, and write the letter of that picture after 2 on the line at the right. The first sample is done correctly. Picture C is **used with** picture 1, so "C" is written after 1 on the line at the right. B is **used with** 2, SO WRITE "B" ON THE LINE AT THE RIGHT AFTER 2. "Nail," marked A, is used with "hammer" marked 3, SO WRITE "A" AFTER 3 ON THE LINE AT THE RIGHT.

NUMBERED

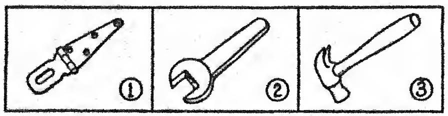
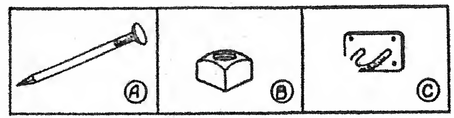


Fig. 1.

LETTERED



(Write answers here)

1. C
2. ___
3. ___

Under each set of pictures you will find some questions. In **each** square at the right of the questions, write a number or a letter to show which tools you would use to do what is asked. Where there are two squares, be sure to write a number or a letter in **each** square. Pictures 3 and A show what is used to fasten a board to a box, so 3 and "A" are written in the squares at the right after question 1. Picture 2 is the correct answer for question 2. Pictures 1 and C are the correct answers for question 3, SO WRITE 1 IN THE FIRST SQUARE AND "C" IN THE SECOND SQUARE AFTER QUESTION 3.

In **each** square at the right of the questions below, write a number or a letter to show which tool in Figure 1 you would use:

(Write an answer in EACH square)

1. To fasten a board to a box.	3	A
2. To tighten a nut.		2
3. To fasten a door so as to use a padlock.		

FIG. 27—An illustrative section from the O'Rourke Mechanical Aptitude Test. These items illustrate the principle but not the difficulty of the test.

the element of familiarity with actual tools or machines. The test consists of a series of items, such as the one illustrated in Figure 28. These items vary from simple problems, which are successfully solved by nearly 100 per cent of industrial applicants, to problems so difficult that an extremely high degree of visualization is required for their solution. An illustration of the use of an early form of this test in a practical situation is given on page 116.

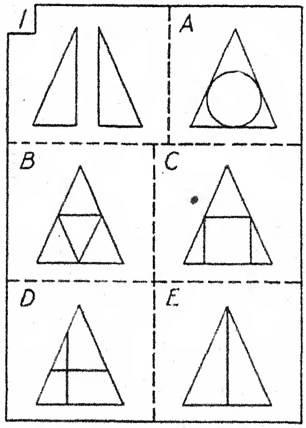


FIG. 28—An item from the Minnesota Paper Form Board Test. The person tested is asked to select the set of lettered parts (*A*, *B*, *C*, *D*, or *E*) which may be formed by the parts shown in the upper-left square.

Purdue Mechanical Adaptability Test[38]

The Purdue Mechanical Adaptability Test, Form A, is designed to aid in identifying men or boys who are mechanically inclined and who, therefore, are most likely to succeed in jobs or in training programs calling for mechanical abilities and interests. The test measures one's experiential background in mechanical, electrical, and related activities. The reason for constructing the test to measure experiential background is that a previous study[39] clearly demonstrated that, other things being equal, boys and men who have profited from previous experiences with *electrical* "gadgets," as demonstrated by knowledge they have assimilated, make significantly better progress in a course in practical electricity than do those who previously have shown little interest in electrical appliances and simple wiring problems.

---

[38] Distributed by the Division of Applied Psychology, Purdue University, Lafayette, Indiana.

[39] C. H. Lawshe, Jr., and G. R. Thornton, "A Test Battery for Identifying Potentially Successful Naval Electrical Trainees," *Journal of Applied Psychology*, XXVII (1943), pp. 399–406.

The questions comprising Form A of the Purdue Mechanical Adaptability Test were selected by statistical methods so as to achieve maximum reliability of the final test and as low a correlation as possible with general intelligence. The resulting correlations were as follows:[40]

Reliability of Purdue Mechanical Adaptability Test = ...84.
Correlation with Intelligence (Adaptability Test Scores) = ...33.

The basic points covered by the Purdue Mechanical Adaptability Test are, in several respects, similar to those covered by several other mechanical aptitude tests, such as the Bennett Test of Mechanical Comprehension and certain parts of the Detroit Mechanical Aptitudes Examination. The Purdue test differs from these other tests primarily by virtue of the fact that in its construction a series of deliberate statistical steps were taken to *eliminate* all questions that correlate with intelligence test scores, with the result that, as mentioned above, the test scores correlate only .33 with intelligence test scores. Most other tests in this field show a much higher relation to intelligence tests, an undesirable circumstance if, as is usually assumed by the particular test user, the test is intended to measure mechanical aptitude as distinct from mental alertness or general intelligence.

Several studies indicating the validity of the Purdue Mechanical Adaptability Test will be summarized on page 119.

The Industrial Training Classification Test[41]

The Industrial Training Classification Test is designed to aid in selecting those persons who are most likely to profit from an industrial training program. It is intended to evaluate an individual's ability to read simple measurements and solve simple arithmetical problems involving whole numbers,

---

[40] C. H. Lawshe, Jr. and Joseph Tiffin, *Preliminary Manual for the Purdue Mechanical Adaptability Test*, (Division of Applied Psychology, Purdue University, Lafayette, Indiana, 1946).

[41] Distributed by Science Research Associates, 228 South Wabash Avenue, Chicago, Illinois.

fractions, decimals, and the conversion of decimals to common fractions. These simple but basic skills are essential to success in many trade schools and industrial training programs.

MacQuarrie Test for Mechanical Ability[42]

Although this scale is called a mechanical ability test, it deals more with dexterity than with an understanding of mechanical principles or familiarity with tools and instruments. A description and discussion of it therefore will be given in the section on dexterity tests rather than in the present treatment of mechanical ability tests.

Typical Individual Tests of Mechanical Ability

All of the preceding tests are of the group variety; that is, they are accomplished with paper and pencil and may be given to a large number of applicants or employees at the same time.

Fortunately, these paper and pencil tests usually will accomplish as much as individual apparatus tests that can be given to only one or two persons at a time and are therefore more costly and time consuming to use.[43] Several individual apparatus tests of mechanical aptitude are available, however, and for some purposes are more serviceable than the group tests.

Cox Mechanical Models Test[44]

This test consists of a series of mechanical models such as the one illustrated in Figure 29. The person tested is shown the view of the model illustrated on the left of this figure and is shown further, by the operation of the model, that as Lever A is moved up and down, Lever B follows this vertical movement. The task of the person being tested is to show, by

---

[42] See page 129.

[43] Willard Harrell, "A Factor Analysis of Mechanical Ability Tests," *Psychometrica*, V (1940), pp. 17–33.

[44] J. R. Cox, *Mechanical Aptitude* (Methuen and Company, Ltd., London, 1928).

simple sketch or design, what mechanism must be involved in the model in order to accomplish this result. The correct mechanism is shown in the sketch at the right, and the sketching of the mechanism is considered the correct solution of the problem. The test in all contains eight such models varying in complexity from the rather simple type of problem illustrated in Figure 29 to problems considerably more complex.

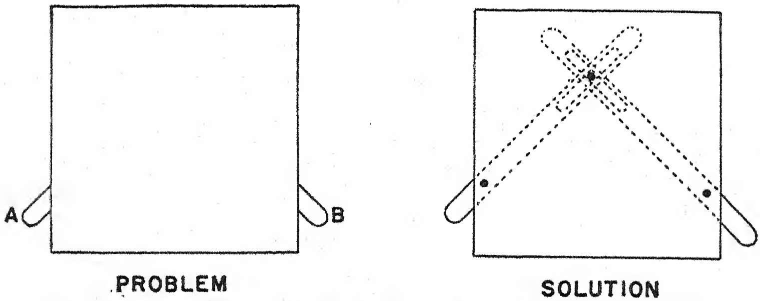


Fig. 29—One of the problems from the Cox Mechanical Models Test.

It should be mentioned that the Cox tests also include a number of printed problems, involving somewhat similar models, which may be given to groups. The Cox tests in their entirety, therefore, embrace both the individual and group forms of testing.

Minnesota Mechanical Assembly Test[45]

This is another individual test of mechanical ability which has been found useful in a number of industrial situations. One form of the test is illustrated in Figure 30. It consists essentially of a number of commonplace mechanical devices, such as a safety razor, a mousetrap, and an electrical plug fixture, all of which may be easily taken apart and re-assembled. The person tested is required to assemble the parts of each article into the finished product. One criticism that

[45] D. G. Paterson, R. M. Elliott, L. D. Anderson, H. A. Toops, and E. Heidbreder, Minnesota Mechanical Ability Tests (University of Minnesota Press, 1930).

has been leveled against this type of test is that the persons tested are likely to vary in their familiarity with the various items, and hence the test is likely to measure this familiarity rather than mechanical ability. In spite of the validity of this criticism from a theoretical viewpoint, the fact remains that, in practical situations, the test has shown itself to be a

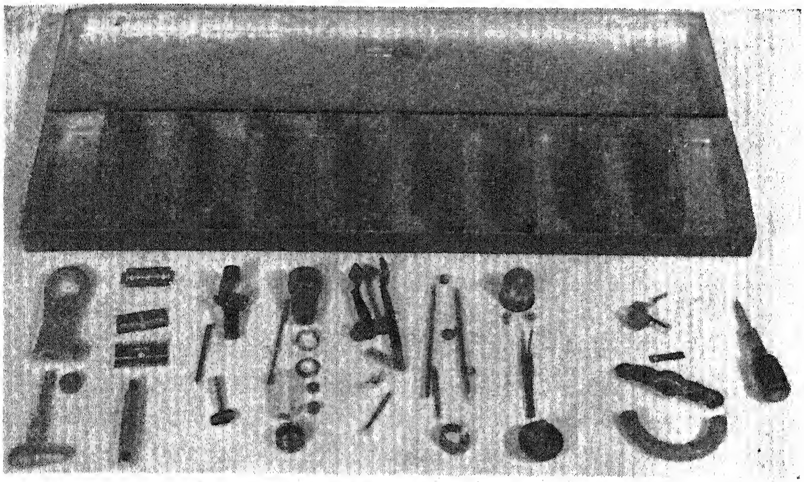


FIG. 30—One form of the Minnesota Assembly Test.

serviceable measuring scale. This fact, rather than a theoretical criticism, is the basis upon which it should be evaluated.

O'Connor Wiggly Block Test[46]

Another test that has become well known within the past few years, particularly among industrial men, is the O'Connor Wiggly Block Test. This test consists of nine pieces of wood cut after the fashion of a three-dimension jig-saw puzzle so that the parts may be assembled into a single rectangular block. Experimental investigations of this test, such as the one reported by Remmers and Schell,[47] indicate that its

[46] Johnson O'Connor, *Born That Way* (The Williams and Wilkins Co., 1928).

[47] H. H. Remmers and J. W. Schell, "Testing the O'Connor Wiggly Block Test," *Personnel Journal*, XII (1933), pp. 155–159.

validity is high within the limit of its reliability, but that the coefficient of reliability of the test is only .35. This low reliability is no doubt due, at least in part, to the fact that the entire test consists of assembling a single block; that is, the test really consists of only a single item or task. Experience has shown that the reliability of a test is nearly always increased by increasing the length or content of the test. This result is due to the simple fact that, with a large sampling of behavior responses of the person tested, a chance success or failure on one or two items will not appreciably affect the total score. But if a whole test consists of only one item or task, the chance element becomes very important in determining the final test score. Thus with only a single test item, such as the Wiggly Block, high reliability could hardly be expected. It is quite possible that the reliability of this test would be much higher if it consisted of a series of such blocks similar in principle but differing slightly in design.

Since such a set of test blocks is not at present available, it is suggested that users of the Wiggly Block Test use it in conjunction with other tests of mechanical comprehension and interpret the Wiggly Block score only as a part of a more complete test profile. This procedure will avoid the mistakes in measurement that are likely to result when complete confidence is placed in a one-item test.

Purdue Mechanical Assembly Test[48]

This test, illustrated in Figure 31, consists of a series of eight boxes of uniform floor area. In each box a mechanism may be assembled so that a certain type of mechanical action takes place. The person tested is first shown the nature of the task by means of a simple illustrative box, and then is allowed a certain predetermined time for assembling the mechanism in each box.

---

[48] M. R. Graney, "The Construction and Validation of a New Type of Mechanical Assembly Test" (A Thesis submitted to the faculty of Purdue University in partial fulfillment of the requirements for the degree of Doctor of Philosophy, 1942).

The test involves several changes over previous mechanical assembly tests. The mechanisms involved are entirely new to all the subjects tested. They do not consist of familiar objects such as have been utilized in many mechanical assembly tests. The series of boxes encompasses every principle of mechanical operation, that is, various types of levers, gears, rack, pinion, worm, and so on. A sufficient

FIG. 31—Purdue Mechanical Assembly Test.

number of assembly tasks are included so that a satisfactory reliability has been achieved ($r = .88$). Studies on the validity of this test are discussed on page 118.

General Versus Specific Factors in Mechanical Ability

In considering the preceding tests of mechanical ability or comprehension, several questions are likely to arise in the mind of the reader. Why is it necessary to have several tests of mechanical ability rather than just one? Is mechanical ability a single, unitary characteristic, like height or weight, or is it a series of partially unrelated abilities, like musical

talent? Are there different types of mechanical ability that are needed in different amounts for success in different mechanical jobs?

Although these questions have not been finally answered by psychological experiments, sufficient data are available to justify certain general statements upon which the answers depend. Evidence from several studies shows that some degree of general or common element is involved in the skills or abilities that we ordinarily think of as mechanical abilities. In Stenquist's[49] early work, it was found that a series of mechanical tests showed average correlations with each other in the neighborhood of .65. Such high intercorrelations can be explained only in terms of a general factor operating in all of the tests. The work of Cox[50] also supports the concept of a general factor in operations in which the subject is called upon to deal mentally with mechanical movements.

In support of the contrary idea—that most mechanical tasks are essentially different from one another and consequently call for a specific type of mechanical ability—Viteles[51] mentions specifically the work of Perrin,[52] Muscio,[53] and Seashore.[54] However, a comparison of the kinds of tests used by Stenquist and Cox, on the one hand, and Perrin, Muscio, and Seashore, on the other, reveals that the tests employed by the former men deal primarily with mental comprehension or understanding of mechanical principles, whereas those used by the latter group are primarily tests of

---

[49] J. L. Stenquist, "Measurements of Mechanical Ability," *Columbia University Contributions to Education* (1923), No. 130.

[50] Cox, *op. cit.*

[51] M. S. Viteles, *Industrial Psychology* (W. W. Norton & Co., 1932).

[52] F. A. C. Perrin, "An Experimental Study of Motor Ability," *Journal of Experimental Psychology*, IV (1921), pp. 24–56.

[53] B. Muscio, "Motor Capacity with Special Reference to Vocational Guidance," *British Journal of Psychology*, XIII (1922), pp. 157–184.

[54] R. H. Seashore, "Stanford Motor Skills Unit," *Psychological Monographs*, XXXIX (1928), pp. 51–66, and "Individual Differences in Motor Skills," *Journal of General Psychology*, III (1930), pp. 38–66.

muscular co-ordination or dexterity. In view of this marked difference in the nature of the tests themselves, it is not surprising that different conclusions concerning the existence of a general factor were reached. In the mental aspects of mechanical ability, that is, in the understanding of mechanical principles, there seems to be considerable evidence for the existence of a general factor, which, when present, will enable the applicant to perform well on most or all tests of mechanical comprehension. But in the manipulative aspects of muscular performance, we most certainly agree with Seashore's conclusion that "the independence of skills measured by these [motor] tests argues against any theory of general motor ability."

A consideration of the skills demanded of the industrial tradesman or skilled machine operator indicates that this employee usually succeeds or fails in proportion to his training and general mechanical comprehension, not in proportion to his basic dexterity. This fact does not mean that successful tradesmen do not need skilled movements, but it does mean that such muscular co-ordination as may be needed can be developed by the majority of tradesmen in training and that it is a lack of mechanical comprehension, rather than inability to develop the muscular aspects of the job, that may prevent them from becoming really proficient in this line of work. This conclusion was also reached by Harrell,[55] who concluded that mechanical ability consists principally in the ability to visualize certain aspects of the job and only to a negligible extent in muscular dexterity. For this reason, tests such as the Bennett, Stenquist, O'Rourke, Revised Minnesota Paper Form Board, and the Purdue Mechanical Adaptability Test are more serviceable in selecting apprentices for the several trades than are tests of muscular co-ordination or dexterity. It is unlikely, however, that any one of the above tests can

[55] W. Harrell, "Testing the Abilities of Textile Workers," *State Engineering Experiment Station Bulletin, Georgia, II* (1940), pp. 14.

be used as a substitute for the others. As mentioned in their respective descriptions, these tests differ markedly in content, and these differences often result in one or another of the tests being adapted to specific industrial jobs. In evaluating these tests for any specific purpose it is necessary to determine experimentally which type of test correlates with known ability on the job in question. Any company contemplating the use of mechanical comprehension tests should carry on sufficient experimentation in its own plant to be sure that the tests used are actually measuring what is desired in the employee. The following sampling of studies of this type illustrates what may be expected of these tests if care is used in their selection and validation.

Uses of Mechanical Ability Tests

Selection of pressmen and machine operators

Hall[56] has reported a study in which the relationship was determined between scores on the Minnesota Paper Form Board Test and ratings of skill of 89 job and cylinder pressman apprentices. A correlation of .58 was found between test scores and rated job ability. It was also found that 70 per cent of the inferior workers obtained test scores below 45, whereas only 6 per cent of the workers of average skill and 5 per cent of those of superior skill received scores below 45. Hall concludes from this investigation that this test may profitably be used as one tool in the selection of pressman apprentices.

Shartle[57] has reported similarly favorable results from a battery of tests designed to select electrical troublemen. Shartle's final battery consisted of six tests, several of which were similar in general content to the mechanical ability tests described in the preceding section. His results show that the

[56] O. M. Hall, "An Aid to the Selection of Pressman Apprentices," *Personnel Journal*, IX (1930), pp. 77–85.

[57] C. L. Shartle, "A Selection Test for Electrical Troublemen," *Personnel Journal*, XI (1932), pp. 177–183.

scores on this test battery correlated .67 with foremen's ratings of ability on the job.

In an unpublished study by the author, a correlation of .47 was found between scores on the Bennett Test of Mechanical Comprehension and foremen's ratings of job performance for a group of forty-seven paper machine operators. A

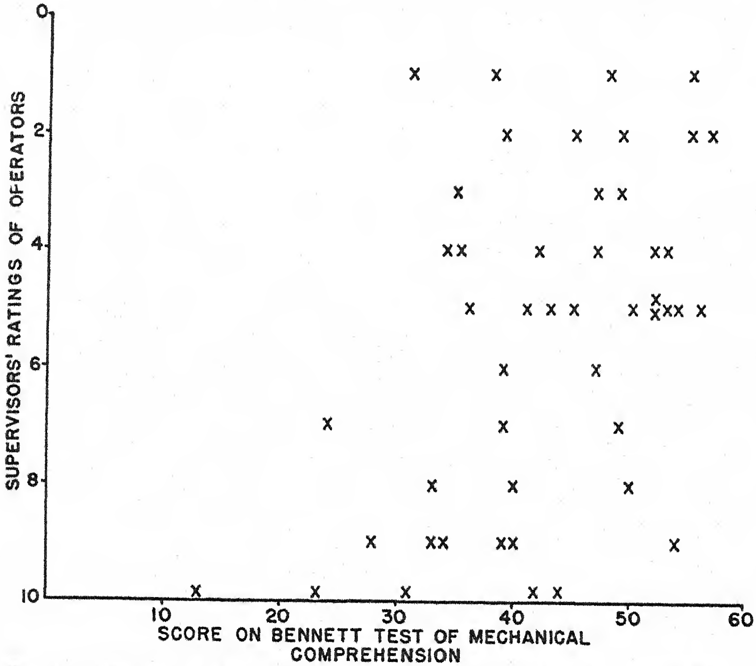


FIG. 32—Relation between score on Bennett Test of Mechanical Comprehension and rated ability of paper machine operators.

scattergram showing the relationship between rated job performance and test scores for this group is shown in Figure 32. Although the relationship as revealed in this scattergram is far from perfect, it will be noted that a critical score set anywhere between thirty and fifty items correct will divide the employees so that those scoring above the critical score will include a definitely higher percentage of high-ability employees than will be found among those scoring below the

critical score. For example, ten of the eleven employees testing above fifty-one on this test are in the top half of the distribution according to rated performance on the job. Harrell[58] reports a similar study of the relationship between mechanical ability test scores and the job performance of cotton mill machine fixers. In this investigation it was found that a fifteen-item adaptation of the Stenquist test showed a correlation of .42 with foremen's ratings of the job performance of the employees.

In a study of power sewing machine operators conducted by Otis,[59] it was found that a battery of tests which included the Minnesota Paper Form Board Test yielded a multiple correlation of .57 against the quality of work of these operators.

An investigation with the Purdue Mechanical Assembly Test[60] showed that this test correlated .55 with supervisors' ratings of machinists and machinists' helpers in one plant. In another plant, a correlation of .35 was found between the test scores and ratings of machinist apprentices by their instructors.

A study using the Purdue Mechanical Adaptability Test[61] on a group of fourteen refrigerator plant mechanics and maintenance men resulted in the scattergram shown in Figure 33. In this study, the rank order correlation between test scores and ratings was .81.

In another study using the Purdue Mechanical Adaptability Test, 46 machine operators from a screw manufacturing company were rated by supervision for their proficiency on the job. A chart showing the percentage of superior operators receiving superior ratings among groups scoring succes-

[58] W. Harrell, "The Validity of Certain Mechanical Ability Tests for Selecting Cotton Mill Machine Fixers," *Journal of Social Psychology*, VIII (1937), pp. 279–282.

[59] J. L. Otis, "The Prediction of Success in Power Sewing Machine Operating," *Journal of Applied Psychology*, XXII (1938), pp. 350–366.

[60] Graney, *op. cit.*

[61] C. H. Lawshe, Jr., Irene Semanek, and Joseph Tiffin, "The Purdue Mechanical Adaptability Test," *Journal of Applied Psychology*, XXX (1946), pp. 442–453.

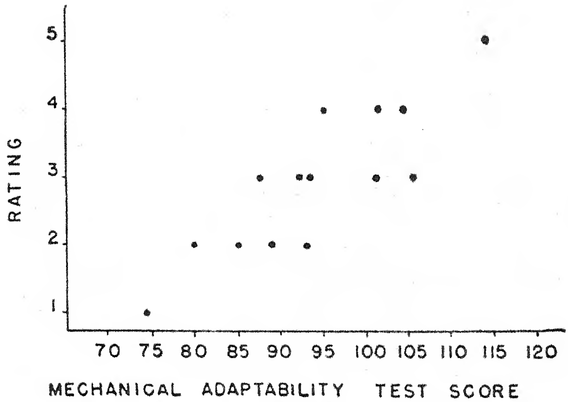


FIG. 33—Relation between scores on Purdue Mechanical Adaptability Test and rated ability of 14 refrigerator plant mechanics and maintenance men.

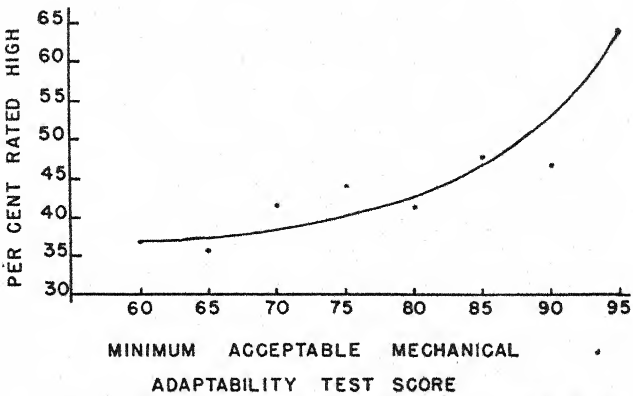


FIG. 34—Percentage of automatic screw machine operators who are rated "high" when successively higher test scores on the Purdue Mechanical Adaptability Test are used.

sively higher on the Purdue Mechanical Adaptability Test is shown in Figure 34. This chart indicates that of men scoring 60 or more (a group that included all the operators), only 37 per cent were rated superior, while among operators scoring 95 or above, 65 per cent were rated superior operators.

McMurry and Johnson[62] report an investigation on the selection of welders in which they found a correlation of .66 between the ability of the men and their scores on the Bennett Test of Mechanical Comprehension. Shuman has reported two studies[63] in which significant correlations were found between ability of both employees and supervisors in an airplane engine and propeller plant and their scores on the Bennett Test and Minnesota Paper Form Board. In scoring both operators and supervisors in these studies, Otis Mental Ability Test scores were also found significantly related to job success.

Selection of students for vocational training

Tests of mechanical ability or comprehension are not limited in their application to the machine operating and maintenance jobs of modern industry. A number of investigations have shown that such tests are ideally suited to be a part of an aptitude test battery for the prediction of success in several types of vocational training. For example, Laycock and Hutcheon[64] have reported that a test battery consisting of high-school grades, the American Council Psychological Examination, the Cox Mechanical Aptitude Test, the Minnesota Paper Form Board Test, and a questionnaire covering interest in physical science gave a multiple correlation of .66 with later achievement in engineering courses.

From such investigations as these we may safely conclude that mechanical aptitude or comprehension tests furnish a real aid to the employment manager in the selection of employees for a wide variety of industrial jobs. In general, these tests do not measure the same capacities or abilities

---

[62] R. N. McMurry and D. L. Johnson, "Development of Instruments for Selecting and Placing Factory Employees," *Advanced Management*, XX (1945), pp. 113–120.

[63] J. T. Shuman, "The Value of Aptitude Tests for Supervisory Workers in the Aircraft Engine and Propeller Industries," *Journal of Applied Psychology*, XXIX (1945), pp. 156–160 and pp. 185–190.

[64] Laycock and Hutcheon, *op. cit.*

that are measured by general mental ability tests, as evidenced by the fact that a study conducted in the Hawthorne works of the Western Electric Plant[65] shows a correlation, for 749 employees, of only .39 between the Otis Self-Administering Test of Mental Ability and the Minnesota Paper Form Board Test. This does not mean that one should adopt the either-or philosophy of using a mental ability test *or* a mechanical comprehension test. Rather it means that one should use in proper combination whatever tests have been found individually to relate to success on the job in question.

What Mechanical Aptitude Tests Should Be Used

The remarks made on page 87 relative to the choice of a suitable intelligence test for any given job apply also to the choice of a mechanical aptitude test.

The several available mechanical aptitude tests differ in content, emphasis, extent to which they measure aptitude or immediate knowledge, and in several other respects. As a result of these differences, the several tests are of quite different value for different jobs of placement. No one can say with certainty which test will be of greatest value in selecting employees for a specific job without carefully studying the relationship between test scores of employees on that job and their actual job success. The fact that several tests of mechanical aptitude are available increases the chances that at least one of these will be suitable for any specified purpose if a sufficient study of the job is made to locate the proper test. With experience, one can make a reasonably judicious choice from a careful study of the job and an analysis of the content of the various tests available. But such a choice should always be checked by correlating the test scores against actual success on the job; and the final decision as to whether or not the test is to be used should be based upon the size of this correlation rather than upon a subjective judgment as to the suitability of the test.

[65] Western Electric Company, *op. cit.*

5

Dexterity, Manipulative, and Achievement Tests

A PREVIOUSLY discussed principle of testing is that usually no single test is sufficient for prediction of job success. Just as most jobs call for a combination of aptitudes, so adequate placement calls for a combination of psychological tests. In addition to determining the mental ability or mechanical aptitude requirements of a job, it is often necessary to determine the amount of muscular co-ordination, bodily dexterity, and aptitude for manipulative work the job requires.

Dexterity Not Related to Mental Ability or Bodily Measurements

Dexterity and mental ability

The measurement of dexterity requires entirely different kinds of tests from those used for measurement of mental ability. The layman often assumes that the person who is "clever with his head" is also "clever with his hands," and that the individual who is high in mental ability is likely to be equally high in bodily dexterity, muscular co-ordination, and aptitude for learning skilled movements. Such is not the case. Numerous investigations have shown that the correlation between muscular skills and mental skills is very close to zero. This circumstance, of course, does not mean that the individual who is above average in one is likely to be below average in the other; but it does mean that it is impossible to predict whether an individual who is high in one will be high,

average, or low, in the other. In an investigation of the Western Electric Company,[1] the correlation between the O'Connor Finger Dexterity Test (described on page 124) and the Otis Mental Ability Test was found to be .07. The correlation between the Hayes Pegboard (another dexterity test described on page 125) and the Otis Test was found to be zero. Each of these correlations was based on test scores of 749 employees. These results are typical of the results of many similar investigations. They prove that one cannot measure finger or hand dexterity by means of a mental test, nor mental ability by means of a dexterity test. If one wishes to determine how much dexterity an applicant possesses, he must measure that dexterity by means of a dexterity test designed for a particular purpose.

Dexterity and bodily measurements

Neither is there any relation between anthropometric measurements and dexterity. Some employment managers judge the dexterity of an applicant by examining his hands and fingers; but when careful anthropometric measurements are made and the results correlated against the measured dexterity of the applicants, no significant relationships have been found. This conclusion was reached by Griffitts[2] and has been corroborated by other investigators. Perhaps in extreme cases, where an applicant has fingers that are stiff or very stubby, one could predict from an examination of his hands that he would probably be low in finger dexterity; but in the great majority of cases such a judgment would be no more than a guess. What an applicant can do with his hands, not the appearance of the hands, determines his qualifications for a manual dexterity job.

---

[1] *Analysis of 1935–37 Experience in Selecting New Men for Shop Occupations.* Privately Printed Monograph. (Western Electric Company, Hawthorne Plant, 1939.)

[2] C. H. Griffitts, "The Relation Between Anthropometric Measures and Manual Dexterity," *Journal of Applied Psychology,* XX (1936), pp. 227–235.

Typical Dexterity Tests

The O'Connor Finger Dexterity Test[3]

This test, illustrated in Figure 35, is a widely used manipulative test. The equipment for giving the Finger Dexterity

FIG. 35—O'Connor Finger Dexterity Test.

Test consists of 310 cylindrical brass pins one inch in length and .072 inch in diameter placed in a shallow tray, about 5 by 6 inches, with gently sloping sides, and a metal plate in which 100 holes have been sunk to a depth of $\frac{3}{4}$ inch with a number

[3] For reference and source of this test, see Table 13 on page 133.

9 drill. The diameter of the holes is .196 inch; they are spaced $\frac{1}{2}$ inch apart, thus forming ten rows of ten holes each.

After a brief period of preliminary instruction the applicant is asked to fill the board—three pins to a hole—as quickly as possible. The frequency distributions shown in Figure 9, page 14, were obtained by means of the O'Connor Finger Dexterity Test. These distributions show that the time required to fill the board varies from around five to fifteen

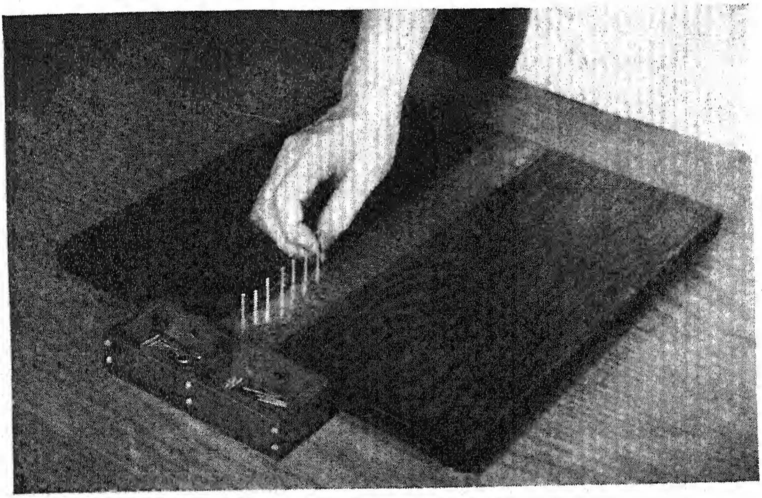


Fig. 36—The Hayes Pegboard.

minutes, depending upon the amount of finger dexterity possessed by the applicant. The reliability of the test, obtained by correlating the first half against the second half, is .98. A number of investigations showing the practical applications of this test in industrial situations are discussed on pages 132–134.

The Hayes Pegboard[4]

This test, illustrated in Figure 36, also measures finger dexterity but involves, in addition, an element of hand and

[4] Elinor G. Hayes, "Selecting Women for Shop Work." *Personnel Journal*, XI (1932), pp. 69–85.

arm co-ordination. The board contains two rows of 26 holes. The rows are one inch apart and the holes in each row are spaced one-half inch apart. Each hole is .125 inch in diameter. The two boxes containing the pins each have inside measurements of $2\frac{3}{8}$ inches by $1\frac{1}{2}$ inches by 1 inch. The pins are $1\frac{1}{2}$ inches long and .115 inch in diameter. The center of the top hole in each row is $\frac{7}{8}$ of an inch from the closest inside edge of the pin box.

The administration of the test calls for a brief (ten or fifteen second) practice period after which the applicant is given nine trials of one-half minute each. These nine trials are arranged as shown in Table 12.

The score on this test is the total number of pins placed during the $4\frac{1}{2}$ minutes of actual testing time distributed in the nine trials summarized in Table 12. Specific studies in which this test has been found of value in placement are discussed on page 136–137.

TABLE 12

SEQUENCE OF TRIALS IN ADMINISTERING THE HAYES PEGBOARD

Trial	*Time*	*Description*
1	$\frac{1}{2}$ minute	Placing pins from right box in right row with right hand.
2	$\frac{1}{2}$ minute	Placing pins from left box in left row with left hand.
3	$\frac{1}{2}$ minute	Simultaneously placing pins from right box in right row with right hand and from left box in left row with left hand.
4	$\frac{1}{2}$ minute	Placing pins from right box in right row with right hand.
5	$\frac{1}{2}$ minute	Placing pins from left box in left row with left hand.
6	$\frac{1}{2}$ minute	Simultaneously placing pins from right box in right row with right hand and from left box in left row with left hand.
7	$\frac{1}{2}$ minute	Placing pins from right box in right row with right hand.
8	$\frac{1}{2}$ minute	Placing pins from left box in left row with left hand.
9	$\frac{1}{2}$ minute	Simultaneously placing pins from right box in right row with right hand and from left box in left row with left hand.

Purdue Pegboard[5]

This test, which is illustrated in Figure 37, measures separately two basic aspects of manipulative dexterity. The

[5] Distributed by Science Research Associates, 228 S. Wabash Ave. For reference, see Table 13, on page 133.

assembly of a series of pin-collar-washer arrangements measures five finger dexterity of the type measured by the O'Connor Finger Dexterity Test. The placing of pins into a series of holes measures manual dexterity of the type involved

FIG. 37—Purdue Pegboard.

in the Hayes Pegboard. These two measurements may be obtained with the same board and require only 2½ minutes of testing time. The Purdue Pegboard results in separate measurements for right hand, left hand, and both hands.

A further advantage of the Purdue Pegboard is that it can be given to ten or more persons simultaneously. All that is required for such group testing is a test board for each person and the careful attention of the test administrator.[6]

[6] In one industrial employment office this test has been successfully given by a single examiner at the rate of fifty applicants per hour. Ten test boards were used simultaneously in this office.

Purdue Hand Precision Test

This test, which deals particularly with precision of hand movement, is illustrated in Figure 38. The applicant punches a stylus successively into holes that are uncovered by a rotating shutter at the rate of 126 holes per minute. The holes uncovered are .5 inch in diameter and are located on the

Fig. 38—Purdue Hand Precision Test.

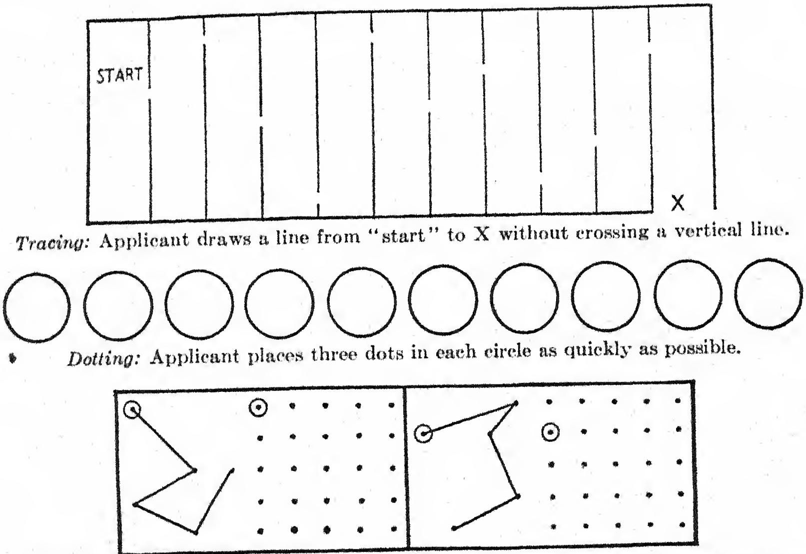
corners of an equilateral triangle measuring 3.5 inches on a side.

The applicant is shown that his task is to punch the stylus into each hole as it is uncovered without allowing the stylus to touch the side of a hole or be caught by the rotating shutter. The shutter is driven by a friction clutch so that in case the stylus is caught the apparatus is not damaged. After a thirty-second practice period, a switch is thrown that connects a cumulative timer which operates whenever the stylus is in

contact with the side of a hole or the shutter. The test consists of a two-minute period of punching which follows immediately and without interruption after the practice period. The score is the number of seconds of contact time (contact of stylus with side of hole or shutter) occurring during the two-minute testing period. The score measures *inaccuracy;* thus the larger the figure, the *poorer* the score. Studies in which this test has "come through" are summarized on page 306.

MacQuarrie Test for Mechanical Ability[7]

The dexterity tests described above are instrumental and all but the Purdue Pegboard must be given individually.



Tracing: Applicant draws a line from "start" to X without crossing a vertical line.

Dotting: Applicant places three dots in each circle as quickly as possible.

Copying: Applicant reproduces the drawings by connecting the appropriate dots in the area to the right of each drawing.

FIG. 39—Several parts of the MacQuarrie Test for Mechanical Ability. This test is essentially a group test of muscular co-ordination and control.

[7] T. W. MacQuarrie, MacQuarrie Test for Mechanical Ability (California Test Bureau, 1925).

This is true of most dexterity tests. A few are of the paper-and-pencil type and are therefore adapted to group testing. Of these, one of the most serviceable is the MacQuarrie Test for Mechanical Ability. Several parts of this test are illustrated in Figure 39. The test measures speed and accuracy of tracing, tapping, dotting, copying, letter location, block identification, and visual pursuit. The first three of the parts deal primarily with manual dexterity. The several parts may be scored separately, which makes it possible to correlate each part against a criterion and determine the optimal weighting of the parts for any specific type of predictive use.

Minnesota Rate of Manipulation Test[8]

This test consists of 58 round blocks $1\frac{1}{16}$ inches in diameter and $\frac{7}{8}$ of an inch thick, and a board approximately 3 feet long and a foot wide containing 58 holes, each $1\frac{3}{4}$ inches in diameter. Two tests may be given with this equipment, namely, the placing test and the turning test. The placing test requires that all the blocks be placed in the holes in the board. The turning test requires that all the blocks be turned over. In addition to the early descriptions of this test, later work on norms and a modified method of administration have been published, respectively, by Cook and Barre[9] and by Jurgensen.[10]

Other dexterity tests

A dexterity test that worked well in placing operators on the job of looping in the hosiery industry makes use of a

---

[8] H. J. Green, I. R. Berman, D. G. Paterson, and M. R. Trabue, "A Manual of Selected Occupational Tests," *Bulletin of the Employment Stabilization Research Institute II*, (1933), pp. 59–85.

[9] D. W. Cook and M. F. Barre, "The Effect of Specialized Industrial Norms on the Minnesota Rate of Manipulation Test as a Selective Instrument in Employment Procedure," *Journal of Applied Psychology*, XXVI (1942), pp. 785–792.

[10] C. E. Jurgensen, "Extension of the Minnesota Rate of Manipulation Test," *Journal of Applied Psychology*, XXVII (1943), pp. 164–169.

grooved pegboard. This test calls for a board containing five rows of five holes each. The holes are .125 inch in diameter and are spaced $\frac{1}{2}$ inch apart. On one side of each hole a groove has been cut. The location of these grooves is varied randomly from one hole to another. Each of the pins to go in the holes has a key along one edge so that the pin may be inserted in the hole only when the key and groove are located correctly with respect to each other. This test is referred to as the Purdue Grooved Pegboard. The test is scored by determining the average time required to fill the board for the third and fourth trials of four that are allowed.

Many other tests of dexterity, muscular co-ordination, or manipulative ability are available. Some of these are specifically adapted to certain jobs and are of little value for other types of placement. Others have been found suitable as placement tests for several different kinds of jobs. Before accepting any one of these tests as suitable for any given job, one should be sure the test actually "comes through" in an evaluation of the type described on pages 52–62. The names and sources of some of the more commonly used tests of this type are summarized in Table 13.

The Uses of Dexterity Tests

It has been amply proved both by experiment and experience that the finger and manual dexterity of applicants, as well as their muscular co-ordination, is revealed much more accurately by dexterity tests than by any other procedure that may be used at the time of placement. As Hurt[11] has said, "Remarkable differences among individuals have been revealed by these tests, differences which a simple interview would never have brought to light." The magnitude of these differences in the case of finger dexterity has already been indicated by Figure 9 on page 14. It remains now to discuss

[11] J. Hurt, "Evaluating Applicants by Dexterity Testing," *Employment Service News*, VI (1939), pp. 7–8.

the relation between dexterity, as measured by such tests, and efficiency on the job after the employees have been placed.

Finger dexterity in watch making

Candee and Blum[12] have reported a study of the relation between finger and tweezer dexterity and the efficiency of workers in a watch factory. They found a statistically significant difference between average scores on the O'Connor Finger Dexterity Test of the superior and mediocre workers. This relationship is represented by a correlation of .26 between foremen's ratings and test scores. The fact that the obtained correlation was no higher was due, in all probability, to the unreliability of the foremen's ratings. However, even with a validity coefficient of .26, the test will have real value as a placement device if a sufficiently small selection ratio is utilized. A tweezer dexterity test[13] (a similar test in which a pair of tweezers instead of the fingers is used in placing the pins), on the other hand, showed practically no relationship with foremen's ratings. One might conclude, *a priori*, that since much of the work in a watch factory is done with a pair of tweezers, a test involving tweezer dexterity would be a more desirable test for this type of work than one involving finger dexterity. The experimental facts, however, support the opposite conclusion. This case will serve to illustrate again the fact that tests should not be accepted merely because of logical considerations in the absence of experiments which clearly show whether they will or will not work for the job in question.

Finger dexterity in small assembly work

A study of the O'Connor Finger Dexterity Test in selecting electrical fixture assemblers and radio assemblers has been

---

[12] B. Candee and M. Blum, "Report of a Study Done in a Watch Factory," *Journal of Applied Psychology*, XXI (1937), pp. 572–582.

[13] For reference and source of this test, see Table 13 on page 133.

TABLE 13

DEXTERITY AND MUSCULAR CO-ORDINATION TESTS

Name of test	What is measured	Source	Reference
Minnesota Rate of Manipulation	Finger and Hand Dexterity	Psychological Corporation, 522 Fifth Ave., New York City	H. J. Green, R. I. Berman, D. G. Paterson and M. R. Trabue, *A Manual of Selected Occupational Tests*, Bulletin of the Employment Stabilization Research Institute II (1933), pp. 59–85.
O'Connor Finger Dexterity..	Finger Dexterity (fine movements)	Human Engineering Laboratory, Stevens Institute of Technology, Hoboken, N. J.	Mildred Hines and Johnson O'Connor, "A Measure of Finger Dexterity," *Personnel Jr.* IV (1926), pp. 379–382.
Purdue Pegboard...........	Finger Dexterity (fine movements) and Finger Dexterity (gross movements)	Science Research Associates, 228 S. Wabash Ave., Chicago, Ill.	Manual of Instructions for the Purdue Dexterity Test, Science Research Associates, 1700 Prairie Ave., Chicago, Ill.
O'Connor Tweezer Dexterity	Dexterity in Using Tweezers	Human Engineering Laboratory, Stevens Institute of Technology, Hoboken, N. J.	Johnson O'Connor, "Born That Way," Williams and Wilkins Co. (1928.
Stanford Motor Skills Unit..	A Variety of Motor Skills	C. H. Stoelting & Co., 424 N. Homan Ave., Chicago, Ill.	R. H. Seashore, "Stanford Motor Skills Unit," *Psychol. Monog.*, XXXIX, No. 2 (1928), pp. 51–66.

reported by Tiffin and Greenly.[14] In the case of electrical fixture assemblers, after experience on the job had been ruled out by partial correlation, the test scores were found to correlate with productivity, as indicated by earnings, to the extent of .22, and to correlate with general efficiency as indicated by merit ratings to the extent of .33. Although these validity coefficients are not so high as might be desired, the statistical chances are ninety-eight out of one hundred that even the lower one represents a real relationship between the test scores and production. If the selection ratio is kept small, say in the neighborhood of .10, one may reasonably expect the test to be of definite value in placing employees on this job.

In the case of radio assemblers, no objective criterion of employee success was available because the wiring of a radio is a line assembly job and the operators are paid on a straight hourly basis. In the absence of differential earnings, pooled ratings of the employees by four raters (the department foreman, line foreman, former line foreman, and personnel manager) were used as the criterion of employee efficiency. These pooled ratings had a reliability of .77. With experience on the job held constant by partial correlation, the correlation between these pooled ratings and finger dexterity test scores was .27. It might be added that a composite test score obtained by combining the finger dexterity scores with scores on the hand precision test (see Figure 38), visual acuity, and color perception gave a multiple correlation of .60 with rated efficiency on the job.

Dexterity on other jobs

An experiment showing the value of dexterity tests in selecting coil winders, operators of punch presses (and similar machines), operators of insulation machines, and bench hands

[14] Joseph Tiffin and R. J. Greenly, "Employee Selection Tests for Electrical Fixture Assemblers and Radio Assemblers," *Journal of Applied Psychology*, XXIII (1939), pp. 240–263.

has been reported by Hayes.[15] Hayes used two dexterity tests—the Hayes Pegboard and the O'Connor Finger Dexterity Test. A scoring system which combined the two test scores, with weights determined by their respective relations to job success, was worked out for each of the jobs studied. The criterion of success in the case of machine operators was output during the first eight weeks on the job. The reliability coefficients for this criterion for the three types of machine operators were as follows:

	Reliability of the Criterion
Coil Winders	.78
Operators of Punch Presses and Similar Machines	.81
Operators of Insulation Machines	.87

The criterion for bench hands was the supervisors' estimates of whether the employees were quick, fair, or slow learners, based on the percentage of standard tasks, which had been set up for the bench-work jobs, attained during the first month. The reliability of this criterion for the four-week period was .89.

The results of Hayes' experiments are summarized in four charts, Figures 40–43. In each of these figures, the key identifies the proportion of employees who were quick, slow, or fair learners, or who did not complete the training. In each case the employees were divided into six groups according to the composite score on the pegboard dexterity tests at the time of hiring. Throughout these results, for all four types of jobs studied, there were, among employees scoring high on the tests, a relatively large proportion of quick learners and a relatively small proportion of learners who were average or slow, or who did not finish the training period. These results are sufficiently clear-cut to justify the use of the dexterity tests as an aid in selecting employees for the jobs studied.

A thorough "follow-up" study of the value of dexterity

[15] Hayes, *op. cit.*

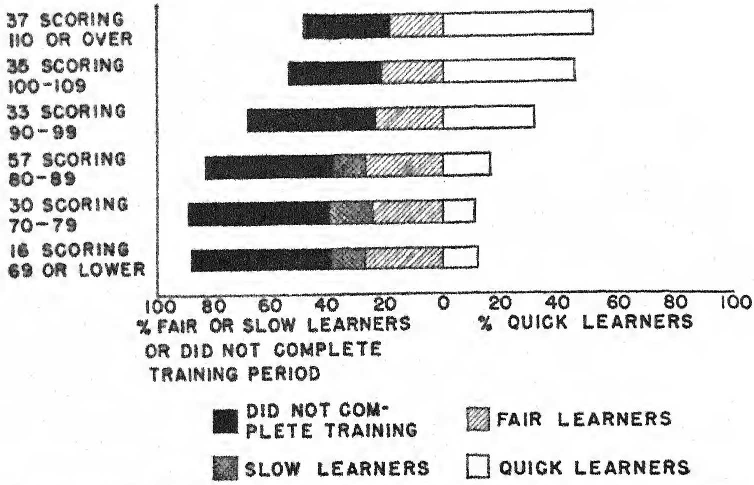


FIG. 40—Relation between composite score on two dexterity tests and speed of learning for a group of 208 coil winders.

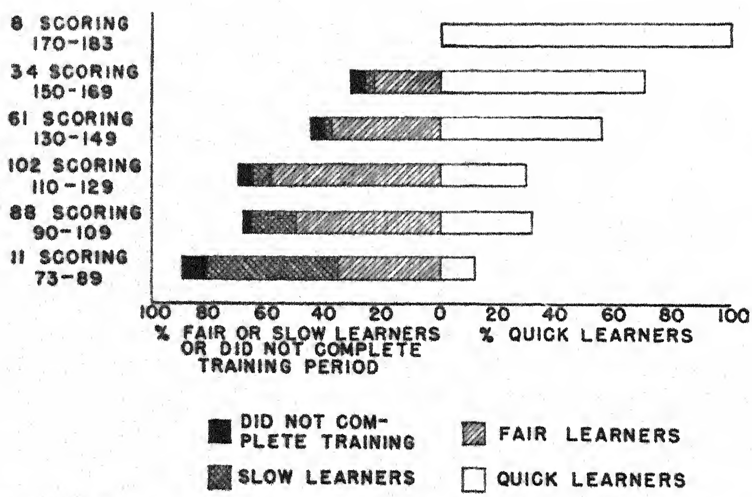


FIG. 41—Relation between composite score on two dexterity tests and speed of learning for a group of 304 bench hands.

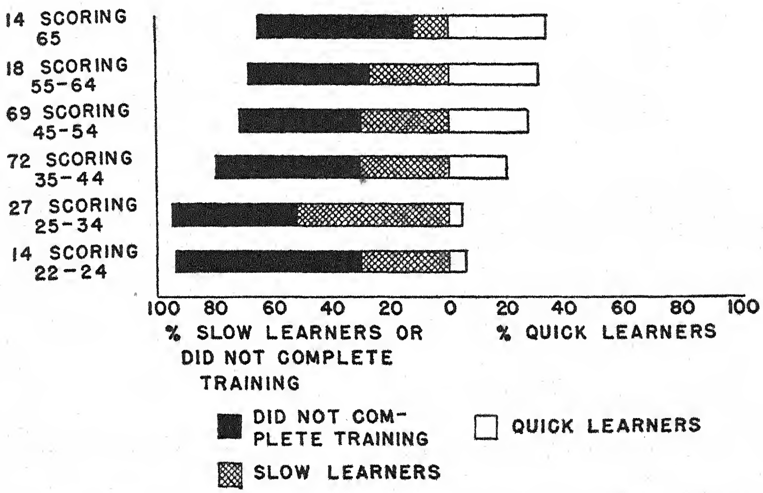


FIG. 42—Relation between composite score on two dexterity tests and speed of learning for a group of 214 insulation machine operators.

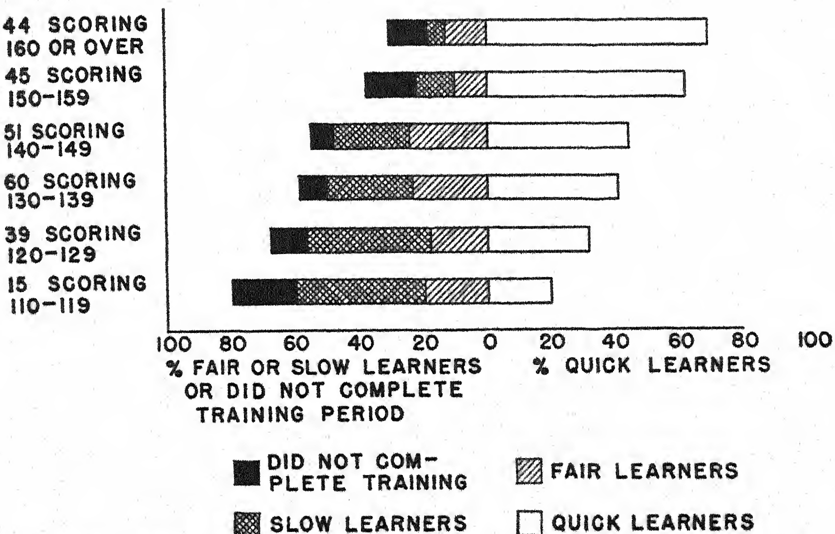


FIG. 43—Relation between composite score on two dexterity tests and speed of learning for a group of 254 operators of punch presses and similar machines.

tests (as well as certain other tests) in the selection of male employees for a variety of occupations has been carried on in the Hawthorne Works of the Western Electric Company.[16] In this study, 749 of the men hired between October 1, 1935, and April 30, 1937, were given a battery of four tests, two of which were the O'Connor Finger Dexterity Test and the Hayes Pegboard. The essential procedure of the study consisted in "follow-up" analyses of the work history of these employees to determine what facts about this work history could have been predicted in a better-than-chance manner from a knowledge of the test scores at the time of hiring.

As there was some variation in the average score on certain tests from one three-month period, or quarter, to another during the six quarters in which the study was conducted, it was necessary to exclude the effect of this time factor in analyzing the results. This was done by separating the records into six quarter-year periods, according to the quarter in which the man entered the company. The analyses were made separately for the men in each quarter and the final result obtained by later combining these results.

Table 14 shows the differences in several criteria of job performance between employees testing above and below average on the dexterity tests. In evaluating the results summarized in this table, it should be kept in mind that the test scores were filed shortly after the employees were hired. The scores could in no way have influenced the judgment of supervisors in determining those criteria of job performance in which judgment is a factor.

In Table 14 the column labeled "high" refers to the employees testing above average on the test. The column labeled "low" refers to employees testing below average. While it will be noted that the high-testing employees are not greatly superior, on the average, to the low-testing employees in the several criteria of job success, it is probably not a matter

[16] Western Electric Company, *op. cit.*

of chance that the high-testing employees on the Hayes Pegboard excel the employees testing below average on this test in six out of the seven criteria.

TABLE 14
DEXTERITY TEST SCORES RELATED TO VARIOUS MEASURES OF JOB PERFORMANCE

CRITERION	O'CONNOR FINGER DEXTERITY TEST			HAYES PEGBOARD		
	High	Low	Diff.*	High	Low	Diff.*
Per cent of persons who have left the company for reasons other than resignation	32.8	43.6	+10.8	34.2	41.6	+7.4
Personnel Rating (1937)	75.9	72.8	+ 3.1	76.1	72.6	+3.5
Increase in weekly earnings since employment (Piece rate employees)	$9.91	$9.54	+ $.37	$10.10	$9.36	+$.74
Increase in weekly earnings since employment (Day rate employees)	$5.24	$5.43	− $.19	$5.54	$5.14	+$.40
Progress in Labor Grade since employment	.76	.78	− .02	.82	.72	+ .10
Per cent of persons without accidents since employment	47.3	45.4	+ 1.9	45.8	47.4	−1.6
Per cent of persons without an illness since employment	88.2	88.3	− 0.1	88.8	87.7	+1.1

* A plus value is given to the difference if it indicates that the high-testing employees are more favorable on the criterion.

The results with the O'Connor Finger Dexterity Test are in general not quite so favorable, but this test seems to be even better than the Hayes Pegboard in selecting employees who are not likely to leave the company for "reasons other than resignation," that is, not likely to be laid off. The difference between the 43.6 per cent of low-testing employees and 32.8 per cent of high-testing employees who left the employ of the company for reasons other than resignation is significant from a statistical viewpoint. One may safely conclude that by using these simple dexterity tests as an aid in placing employees, persons will be placed who will tend to stay longer on the

job, who will be considered the best employees by their superiors, who will show more rapid increases in piece-rate or daily earnings, who will progress more rapidly in labor grade, and who will be at least no more likely to experience accidents or sickness than persons placed without the tests.

A number of other organizations have found tests of dexterity and muscular co-ordination helpful in the placing of employees on certain types of jobs. Typical of such investigations is a series of studies by Drake[17] in the selection of inspectors for the Eagle Pencil Company. Most of the tests that Drake describes are of the manipulative type, though some deal with other aspects of employee aptitude, such as visual perception and acuity. One new pegboard test described by Drake has a reliability of .92 and a validity of .59. A test having a validity coefficient as high as .59 is of considerable value even when it is not possible to work with a low-selection ratio.

Blum and Candee[18] have reported an investigation dealing with the relation between rated job success of packers and wrappers in a department store and scores made on the Minnesota Placing and Turning Test. They found the ratings to be correlated .35 and .27, respectively, with the placing and turning scores on the test.

Blum[19] has published a study in which both production records and supervisor's ratings of sewing machine operators were correlated with scores made on the MacQuarrie Test for Mechanical Ability and an ingenious test, similar to a work sample, known as the Blum Sewing Machine Test. He found that parts of the MacQuarrie Test were significantly related to production records, and that the Blum Sewing

---

[17] C. A. Drake, "Aptitude Tests Help You Hire," *Factory Management and Maintenance*, XCV (1937), pp. 55–57.

[18] M. Blum and B. Candee, "The Selection of Department Store Packers and Wrappers with the Aid of Certain Psychological Tests," *Journal of Applied Psychology*, XXV (1941), pp. 76–85.

[19] M. L. Blum, "Selection of Sewing Machine Operators," *Journal of Applied Psychology*, XXVII (1943), pp. 35–40.

Machine Test (which is described in his report) is also significantly related to job success.

Another illustration of the practical effectiveness of a dexterity test in selecting employees is the case discussed in an earlier chapter (page 79) of loopers in a hosiery mill. It will be remembered that a dexterity test alone provided the means of markedly reducing the learning cost of employees on this job.

An unpublished study on the use of the Purdue Pegboard with employees on an operation known as quilling in a textile

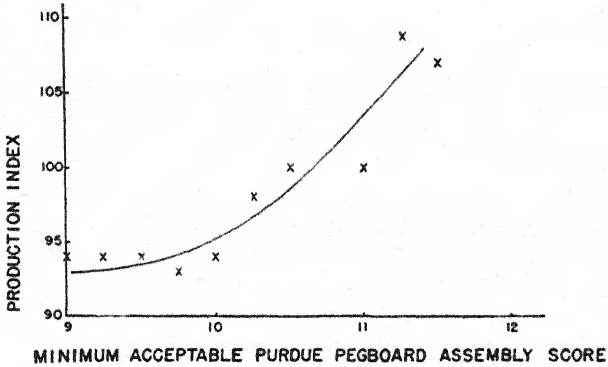


FIG. 44—Job performance of "quillers" in relation to successively higher scores on the assembly part of the Purdue Pegboard. (From data supplied by courtesy of Burlington Mills Corporation, Burlington, N. C.)

mill resulted in the curve plotted in Figure 44. Production indices of twenty-eight operators of approximately equal experience were obtained. In Figure 44, these production indices are plotted against minimum assembly scores on the Purdue Pegboard. As the curve shows, among the operators scoring nine or above (a category including all the operators) the average production index was 94; with successively higher minimum scores, the average production index was also successively higher.

In using dexterity tests, it is particularly important to keep in mind the fact that a test that at first sight *seems* ideal

for a given job may not "come through" in practice; while one that does not *seem* to be so satisfactory may produce results much more highly related to success on the job. This need for caution in the selection of tests is amply proved by such reports as those of Ghiselli[20] on experimental work done with package wrappers and inspector-packers.

Selection of apprentices

Trade apprentices need much more than capacity for developing the manipulative aspects of their trade. They need, first of all, capacity for learning the fund of knowledge that is required of an expert tradesman. This capacity is better measured by tests of mental ability and mechanical comprehension than by tests of dexterity and muscular co-ordination. Yet, in spite of the fundamental importance of these mental capacities, the element of dexterity is also of some importance in achieving success as a trade apprentice. Allen,[21] for example, reports that a battery of seven tests covering intelligence, mechanical aptitude, mechanical ability, and dexterity was found of value in selecting boys who are able to make satisfactory progress in learning the various skilled trades. While three of the four elements measured in this study deal with the mental aspects of the trainees, it is significant to note that dexterity also was of some importance. In a later study, Allen and Smith[22] report that the same battery of tests showed a correlation of .82 between scores obtained before training and after training. The significance of this finding is that such tests may be used with confidence *before* training is given; that is, the ranking of applicants at the

[20] E. E. Ghiselli, "Tests for the Selection of Inspector-Packers," *Journal of Applied Psychology*, XXVI (1942), pp. 468–476, and "The Use of the Minnesota Rate of Manipulation and the O'Connor Finger Dexterity Tests in the Selection of Package Wrappers," *Journal of Applied Psychology*, XXVII (1943), pp. 33–34.

[21] E. P. Allen, "The Selection of Engineering Apprentices," *Journal of the National Institute of Industrial Psychology*, V (1931), pp. 379–384.

[22] E. P. Allen and P. Smith, "The Selection of Engineering Apprentices II," *Human Factor* (London), IX (1935), pp. 63–67.

time of hiring will not be significantly modified by training on the job.

The preceding discussion has been largely concerned with dexterity tests alone. In actual practice, the most effective use of any program of tests usually involves a combination of several types of tests, as in the work of Allen and Smith just described. Typical of such combinations or batteries is the group of tests proposed by Otis[23] for the selection of power sewing machine operators. Otis followed the customary procedure of correlating a large number of tests, each individually against a criterion of job success, and then determining which combination of the original tests best predicted the criterion. He used separate criteria for quality of work and quantity of work, and found that a somewhat different battery should be used in predicting these two phases of job performance. A battery consisting of the Minnesota Vocational Test for Clerical Workers[24] (name comparison), Poppelreuter Tracing Test[25] (time score), the Poppelreuter Weaving Test[26] (paper folding test), the Minnesota Spatial Relations[27] (time score), and Minnesota Paper Form Board Test,[28] Forms A and B, gave a multiple correlation of .57 with the quality of work of the sewing machine operators. In predicting the quantity of work produced by these operators, he found that the O'Connor Tweezer Dexterity Test, the Poppelreuter Tracing Test (time score), the O'Connor Finger Dexterity Test, the Minnesota Rate of Manipulation Test, and the Minnesota Vocational Test for Clerical Workers

---

[23] J. L. Otis, "The Prediction of Success in Power Sewing Machine Operating," *Journal of Applied Psychology*, XXII (1938), pp. 350–366.

[24] Minnesota Vocational Test for Clerical Workers (Psychological Corporation, 522 Fifth Ave., New York, N. Y., 1933).

[25] W. Poppelreuter, "Die Arbeitskurve in der Eignungsprüfung," *Industrielle Psychotechnik*, III (1926), pp. 161–167.

[26] *Ibid.*

[27] D. G. Paterson, R. M. Elliott, L. D. Anderson, H. A. Toops, and E. Heidbreder, Minnesota Mechanical Ability Tests (University of Minnesota Press, 1930).

[28] See page 105.

(number comparison) gave a multiple correlation of .64 with the actual amount of work produced by the employees. It is interesting to note that the Tweezer and Finger Dexterity Tests are both of value in selecting the most rapid sewing machine operators, but are of no help in selecting the individuals who will do the highest quality of work. The use of separate criteria for speed and quality of work, or even of a still larger number of criteria, such as those discussed on pages 53–55, in the validation of tests and batteries of tests, is often of exceptional importance to management. For some kinds of work or some orders, quantity is of greater importance than quality; and for other types of work an opposite situation may prevail. By knowing specifically whether the employees placed by means of any test program will excel in quantity or quality, safety, or other factors, management can decide more easily whether the battery of tests will select employees who will be able to achieve the objectives that have been set up as desirable by company policy. Thus, one company which specializes in an exceptionally high-quality product might use a certain battery of tests in selecting and placing employees, while another company which makes the same product in quantity at a very low price might use quite a different test battery. This example should serve to emphasize again the fact that a test battery should be adopted not only in accordance with the particular job which the employees are doing, but also in accordance with the major policies and objectives of the company. Loopers in a hosiery mill that produces inexpensive hose might be selected by quite a different battery of tests from those used on loopers in a mill that specializes in a high-quality product. The first organization wants operators who are rapid, even though the quality of their work may at times be somewhat imperfect. The second organization wants employees who produce a perfect product even though their output may fall far below that of employees on a similar job in the first-mentioned mill.

Achievement Tests

The tests discussed in the preceding chapter and in the first part of this chapter are essentially tests of aptitude, not of achievement.[29] They are, therefore, of value primarily in the selection of individuals to be trained for various types of work. However, when applicants are being hired who have had or who claim to have had a certain type of work experience, the most serviceable type of selection device is usually an achievement test. For example: suppose that an industry is hiring a number of employees to operate a certain type of machine. The correct operation of the machine requires a certain amount of experience and training. Among the applicants are a few who claim to have had the necessary experience with some other company. Perhaps the best way to determine whether such applicants really are expert machine operators would be to put them on the machine in question for a trial period, give them little or no specific instruction, and determine by try-out on the job how well they can operate the machine. Since often the operation of such a machine is a definite hazard to an individual who is not actually proficient, it would be unwise, and might even be dangerous, to allow an uninstructed applicant to operate an actual machine. It is possible, and in times of depression even probable, that some applicants may profess to have had certain experience that they have not had or may claim a degree of proficiency that they do not possess. It is, therefore, highly desirable to make use of appropriate achievement tests in the selection of such employees. Such tests are simply means of determining in a convenient, simple, and economical manner, whether applicants actually possess the job proficiency that they claim to have.

Job achievement tests may be of several forms. They may consist of standardized oral questions, objective written

[29] The distinction between these two types of tests is made on page 50.

questions of the true-false or multiple choice variety, or, in some cases, of a miniature (and safe) replica of the job itself. Some job achievement tests must be given individually, while others may be readily administered to large groups of applicants simultaneously. The essential characteristic of the achievement test as such, in contrast with the aptitude test, is that the former measures how much actual job proficiency the applicant is able to demonstrate at the time he seeks employment. Achievement tests are therefore of greatest value in the selection of employees for jobs that are reasonably consistent from one industrial plant to another.

Oral Trade Questions

Oral trade questions are among the most widely used types of achievement test. Such questions are convenient to administer and simple to interpret. The most extensive research with this type of test has been carried on by Stead, Shartle, and associates.[30] The general procedure followed by this group has been described in detail in their book, *Occupational Counseling Techniques*. Since the many trade questions developed are in constant use by the various state employment offices, and since the publication of the actual questions used would cause them to lose much of their value, Stead and Shartle publish only a few sample questions to illustrate the method by which the validity of the questions has been determined. This procedure is sufficiently important to the general subject of test validation to warrant a somewhat detailed discussion.

Validation of oral trade questions

The general procedure in validating these questions consists in finding for each trade a few questions, usually fifteen, that can be answered correctly by a large proportion of successful journeymen in that trade, but that are answered

[30] W. H. Stead, C. L. Shartle, and Associates, *Occupational Counseling Techniques* (American Book Company, 1940).

correctly by a definitely smaller percentage of apprentices in this trade, and are answered correctly by a still smaller, or negligible, percentage of employees in related trades. Each question to be retained must show a significant dropping off in the percentage of correct answers from the journeyman group to the apprentice group to the related trades group. When a set of questions for a given trade has been obtained in this manner, it is a simple matter to determine whether an applicant who claims to possess journeyman abilities for that trade actually is able to answer these questions as well as known tradesmen have answered them. The differentiation obtained by such a set of oral trade questions for asbestos workers is shown in Figure 45. The expert asbestos workers vary over a range of seven to fifteen correct answers; apprentices and helpers, of zero to twelve; and related workers, of zero to five. Thus, an applicant who is able to answer correctly only six or fewer of the fifteen questions is very unlikely to be an expert asbestos worker, regardless of what he may say or feel about his own skill. Even if he answers up to twelve items correctly, he is probably not highly proficient in his trade, because the majority of men known to be experts were able to answer thirteen or more of the questions correctly.

Factors affecting the validity of trade questions

The success with which such a set of trade questions will differentiate among applicants possessing different degrees of skill is dependent upon the very careful selection of questions comprising the test. This selection can only be made by starting with many more questions than will finally be retained and by eliminating every question that does not actually show a differentiation among the several groups tested. Likewise, if the test is to be used on a country-wide basis, the preliminary validation work must be carried on in the several geographical areas where the test is later expected to function. The following question for a roofer illustrates the necessity

FORM I

Score	Expert Asbestos Workers (50 Subjects)	Apprentices and Helpers (25 Subjects)	Related Workers (25 Subjects)
15	xxxxxxxx		
14	xxxxxxxxxxxxxxxxx*		
13	xxxxxxxxxxxx		
12	xxxxxx	xx	
11	xx		
10	x	x	
9	x		
8	xx	xx	
7	x	x	
6		xxxxxxxx*	
5		xxx	xxxx
4		xxxx	xx
3		xx	xxx
2		x	xx
1		x	xxxx*
0			xxxxxxxxxx

FORM II

Score	Expert Asbestos Workers (50 Subjects)	Apprentices and Helpers (25 Subjects)	Related Workers (25 Subjects)
15	x		
14	xx		
13	xxxxxxxxx		
12	xxxxxxxxxx	x	
11	xxxxxxxxx*		
10	xxxxxxxx		x
9	xxxxx	x	x
8	xxx	x	
7		xxxx	xxx
6	xxx	xxxx	
5		xx*	x
4		xx	xxx
3		xxxx	xxx
2		xx	xxxx*
1		xxxx	xxx
0			xxxxxx

*** Median Score**

FIG. 45—Distributions of scores for expert asbestos workers, apprentices and helpers, and related workers on a 15-question oral trade test.

for considering geographical factors in the validation of the question.

Q. What type of asphalt is glued on a flat roof?
A. Flat (F) (low melt).[31]

Experts west of the Mississippi River answered this question with unusual consistency, but there tended to be a reduction of consistent answers by experts when the validation took place eastward. Therefore this question could not be used in a test that was intended for country-wide use. In general, questions that could be answered simply, preferably by one or two words, were found to be better than those called for more extensive answers. Often the best questions were those that could be correctly answered only by terms peculiar to the trade in question, even though such terms might not be found in a dictionary.

TABLE 15
TRADE QUESTIONS IN TRUE-FALSE

FORM COVERING THE JOB OF HEATER IN A STEEL MILL

1. The furnace checkers are always located directly below the car bottom.	T	F
2. The gas flow and air flow must always come from opposite sides of the furnace.	T	F
3. The only way to tell if the stock gas is burning is to look at the top of the stock.	T	F
4. The purpose of the stock gas is to help the furnace draw.	T	F
5. In lighting a cold furnace, the rear parts are always the first ones to be opened.	T	F

Trade questions of this general type have been prepared for most of the standard trades and are in use in many state employment offices. The procedure followed has given such excellent results that a number of industries are now adopting this method in the formulation of questions for the use of the employment manager. A number of industrial plants now make up their own questions for their own particular jobs. This procedure enables the test constructor to make up a test that is of much greater value to the industry in question than

[31] *Ibid.*, p. 38.

a test that has been constructed for more general application. Examples of such questions in true-false form covering the job of heater in a steel mill are given in Table 15.

The Miniature or Job Sample

The miniature or job sample method of constructing an achievement test consists in trying out the applicant in a test situation that reproduces all, or an important sampling of, the actual operations that the job itself requires. The miniature test usually consists of apparatus, but the apparatus is constructed to eliminate whatever hazard might be involved in operating a production machine. The applicant is asked to operate the miniature under conditions that simulate the operation of the real machine. Norms of performance of experienced and inexperienced operators are obtained in the test situation so that the status of an applicant can quickly be determined by a comparison of his score with the norms.

A miniature punch press

An example of this approach is the miniature punch press described by Tiffin and Greenly.[32] This test is used in the selection of punch press operators. This apparatus, which is illustrated in Figure 46, is a replica in all essential features of a small industrial punch press. It differs from a real press in that the punch is located in a vertical bearing and is held down only by a spring. This feature prevents the punch itself from descending when an obstacle is encountered. When this occurs, the punch remains stationary while a mechanical counter records an error or mispunch. The test is administered by having an applicant put through the press two hundred pre-punched pieces of galvanized sheet iron. The time required to feed these pieces is recorded by means of a stop watch. During the test period, the mechanical

[32] Joseph Tiffin and R. J. Greenly, "Experiments in the Operation of a Punch Press," *Journal of Applied Psychology*, XXIII (1939), pp. 450–460.

FIG. 46—Applicant being given the Miniature Punch Press Test.

counter records the number of errors or mispunches. The test thus results in a simultaneous time and error score.

The validity of this test has been studied by comparing the average performance of different groups of persons to whom

the test had been given. The curves in Figure 47 summarize the results obtained for three specific groups. These curves show the relation between mispunches or errors in punching two hundred plates and the time in minutes required to punch the two hundred plates. As might be expected, the errors decrease as the time increases for all three groups of persons upon whom these results are based. It is interesting

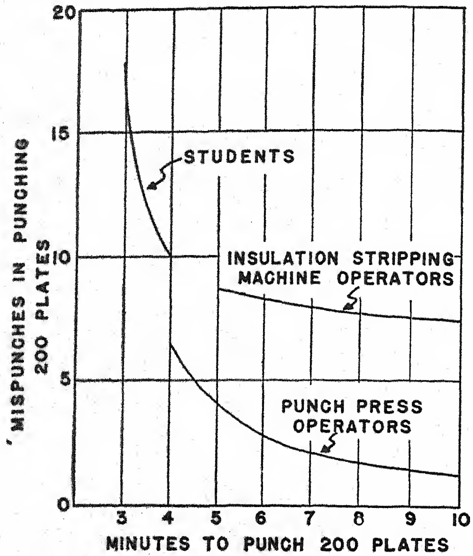


Fig. 47—Relation between speed and accuracy for three groups of subjects in operating a small punch press.

to note, however, that the curve for the students and the one for the insulation stripping machine operators are almost identical mathematically and simply represent different segments of the same curve. This suggests that the students, though punching much more rapidly and hence making many more errors, are not significantly different from the insulation stripping machine operators in genuine ability. The employees in this latter group, though offering no experience in the operation of a punch press, were industrial employees and, as

such, were more accustomed to the need for careful, slow operation of any machine. The curve for the industrial punch press operators, however, is markedly different from that of either the students or insulation stripping machine operators. For any given speed of operation the punch press operators were more accurate than either of the other groups, and for any given level of accuracy, punch press operators were more rapid.

From a knowledge of the speed and accuracy of a given applicant and a comparison of this information with the data graphed in Figure 47, it is possible to determine the status of the applicant in comparison with the corresponding test performance of persons who are known to be experts on this job.

The method illustrated by the miniature punch press has been utilized in selecting employees for numerous other jobs. For example, in hiring persons for such jobs as packaging, inspecting, and certain types of machine operation, it is often extremely helpful for the personnel manager to obtain a sample of the quality and speed of work that the applicant is able to perform. The miniature or job replica furnishes such a sample of job performance with a minimum of effort and with no danger to the applicant while the test is being administered.

The job sample in clerical work

Another area in which the job sample method of testing has resulted in very satisfactory results is the selection of employees for stenographic, clerical, and secretarial positions. One of the most widely used tests of this type is the series known as the Blackstone Stenographic Proficiency Tests.[33] This series includes a stenography test and a typewriting test. The stenography test is made up of seven parts, namely, knowledge of English grammar (which includes punctuation, capitalization, and spelling); syllabication; office practice;

[33] Blackstone Stenographic Proficiency Tests (World Book Company, 1932).

alphabetizing; abbreviations; knowledge of business organization; and ability to take dictation and shorthand and to make the correct transcription. Separate norms are available for each of these seven parts so that in hiring an applicant for a given job it is possible to match the test results with a job analysis of the work to be performed. The second part of the Blackstone series consists of a typewriting test. This part requires the applicant to copy a standard page of typewritten material. The result may be scored separately for speed and accuracy and the norms available permit a rapid determination of the exact level of proficiency possessed by an applicant.

Another standard test used in measuring proficiency in the use of a typewriter is the Thurstone Examination in Typing.[34] A part of this test consists of a page of corrected copy that the applicant is asked to type, making the indicated corrections. A sample of the material to be copied is shown in Figure 48.

No P Business system was used in ascertaining the amount~~s~~ of cir-
culation of various publications as well as the kind~~s~~ of circulation. Advertisements
were keyed, and ~~many~~ other means were employed to discover ~~determine~~ ~~what~~
~~was~~ the exact value of each ~~certain~~ style~~s~~ of advertisement~~s~~ and of each ~~what~~
~~was the best~~ medium in which were ~~to~~ insert ed. ~~them.~~
advertisements

FIG. 48—Part of Test I of the Thurstone Typing Test. The person tested is required to copy this material, making the indicated corrections.

The ability to copy corrected material of this kind, though frequently called for in many stenographic positions, seems to be quite different from the ability to make an exact copy of material requiring no corrections. For example, in one employment office a girl who had recently won a state contest for speed in typing was unable, when tested with the Thurstone test, to obtain even an average score in typing this corrected copy. Apparently her ability consisted of a strictly

[34] Thurstone Employment Tests, Examination in Typing (World Book Company, 1922).

mechanical and automatic method of "copying the copy." While such ability might be very useful in winning state contests, it is of little use in a business situation because one is seldom asked to copy material that is perfect at the outset. A test of the Thurstone type, therefore, is a much better measure of the kind of ability called for in a business or industrial job than is a test that requires only straight copying of material.

Many other standardized tests in the general field of clerical aptitude and achievement are available commercially. A list of some of those in most common use is given in Appendix C.

Written Achievement Tests

The Purdue Vocational Tests[35] comprise a series of paper-and-pencil tests that measure achievement in technical information related to various areas of trade training. These tests, although developed primarily for the use of vocational teachers in public-school systems, are being used also by industrial personnel men for the purpose of determining the qualifications of applicants for jobs connected with the various trades. Several standardized tests are now available in this series.

Technical information in machine shop[36]

This test consists of 133 questions (partly multiple choice and partly matching) that cover material well known to any qualified machinist and that should be reasonably familiar to any boy who has received adequate vocational instruction in this field. Four examples of the multiple-choice items follow:

---

[35] Distributed by Science Research Associates, 228 S. Wabash Ave., Chicago, Ill.

[36] H. F. Owen, C. C. Stevason, H. G. McComb, and C. D. Hume, Technical Information Test for Machinists and Machine Operators (Science Research Associates, 228 S. Wabash Ave., Chicago, Ill.)

In each of the multiple choice statements listed below there are four possible answers, but only one is correct. Read each statement carefully before making your choice of answers.

By tolerance in machine work is meant (*a*) the allowance, as for the oil film of a bearing, (*b*) the amount of variation either above or below a certain basic measurement that will be acceptable, (*c*) the amount of stock left for polishing, (*d*) the amount of stock left for grinding........... (   )

If it fits, the wrench which will probably do the least harm to the corners of a nut is (*a*) an adjustable, (*b*) an alligator, (*c*) an open end, (*d*) a socket wrench.................. (   )

One of the best bearing metals contains antimony, tin, and copper. This metal is called (*a*) bronze, (*b*) brass, (*c*) babbitt, (*d*) lead.................................... (   )

Standard taper sizes are designated by (*a*) decimals, (*b*) fractions, (*c*) numbers, (*d*) letters........................ (   )

The test covers such topics as hand tools, bench tools, bench work, lathe, milling machine, shaper, planer, and drill press. The test is arranged so that subscores may be readily obtained for bench operations and for the operation of the grinder, planer, lathe, and milling machine. On the basis of odd *vs.* even item correlations, the reliability of the various scores was as follows:[37]

Bench operations	.87
Grinder	.90
Planer	.89
Lathe	.80
Milling Machine	.85
Total Score	.96

The published norms were obtained on the basis of a fifty-minute time limit. The scoring is simple and rapid. Separate norms are available for students or applicants offering different amounts of vocational training or practical experience.

[37] Joseph Tiffin, Preliminary Manual for the Purdue Test for Machinists and Machine Operators (Science Research Associates, 228 S. Wabash Ave., Chicago, Ill.).

Blueprint reading test[38]

This test consists of seventy-nine questions covering basic principles involved in reading and interpreting blueprints.

Technical information in electricity[39]

Another test in the Purdue Vocational Series deals with industrial electricity. This test is available in two forms, A and B, each consisting of sixty-five multiple-choice and matching type questions covering material of a practical nature in the field of electricity. Four sample questions from this test follow:

In the following section of this test select the word or phrase after each incomplete sentence which makes the most nearly true statement. Record the number of the choice made.

A bathroom fixture should never use a brass shell socket because it may (1) short circuit and blow fuses, (2) draw excessively high currents, (3) give one a dangerous shock, (4) become corroded and open circuit.................. (   )

Conductors of electricity are most often made of (1) iron, (2) brass, (3) lead, (4) copper........................ (   )

An electric bell depends for its operation upon (1) the heating effect of electricity, (2) the mechanical effect of electricity, (3) the magnetic effect of current flow, (4) the impedance of the circuit...................................... (   )

An accidental connection between two wires of opposite polarity is called (1) a short circuit, (2) a parallel circuit, (3) a series circuit, (4) a ground...................... (   )

This test covers such topics as common electrical circuits, measuring units, conductors, and common electrical devices. Its reliability, computed by the method of chance halves, is .91. The time limit is fifty minutes.

---

[38] H. G. Owen and J. N. Arnold, Purdue Blueprint Reading Test (Science Research Associates, 228 S. Wabash Ave., Chicago, Ill.).

[39] C. W. Caldwell, H. R. Goppert, H. G. McComb, and W. B. Hill, Technical Information Test for Electricians (Science Research Associates, 228 S. Wabash Ave., Chicago, Ill.).

Technical information in industrial mathematics[40]

In its construction, scoring, and administration, this test is similar to the tests just described. It deals primarily with operations of an arithmetical, or simple mathematical, type that a tradesman in industry is likely to encounter. The test is available in two matched forms, A and B. A preliminary computation, based on a group with a rather narrow range of scores, gave a correlation of .82 between the two forms. A sample from this test reads as follows:

This test consists of mathematical problems that are found in various kinds of industrial work. Each problem has FOUR suggested answers, but only ONE is correct. Write an X over the correct answer.

1. The difference between $\frac{1}{3}$ and $\frac{1}{4}$ inch is:..... (1) (2) (3) (4)
   (1) $\frac{1}{6}''$ (2) $\frac{1}{8}''$ (3) $\frac{1}{10}''$ (4) $\frac{1}{12}''$
2. If 28 man-hours are required to build 4 tables, how many man-hours will be needed to build 5 tables?.............................. (1) (2) (3) (4)
   (1) 32 (2) 35 (3) 38 (4) 41
3. If a piston has an area of 6 square inches, and a force of 30 pounds is required to move the piston, the pressure per square inch is:..... (1) (2) (3) (4)
   (1) 5 (2) 6 (3) 7 (4) 8
4. If $3a + 7 = 22$, then "a" equals:......... (1) (2) (3) (4)
   (1) 3 (2) 4 (3) 5 (4) 6

Another series of simplified achievement tests has been published under the series title *Interview Aids*.[41] These include "Can You Read a Micrometer?" "Can You Read a Scale?" and "Can You Read a Working Drawing?"

Achievement tests of this sort in the different trade areas enable an employment manager to obtain a thorough picture of the strong and weak areas of information possessed by applicants. Because the tests may be given to many appli-

---

[40] C. H. Lawshe, Jr. and Dennis H. Price, Purdue Industrial Mathematics Test (Division of Applied Psychology, Purdue University, Lafayette, Ind.).

[41] Distributed by Science Research Associates, 228 S. Wabash Ave., Chicago, Ill.

cants simultaneously, they can be made much longer than oral trade tests and therefore are much more exhaustive than the latter in covering the technical field.

Uses of Written Achievement Tests

Selection of apprentices

Because of the increasing availability of public-school vocational courses, applicants for industrial apprenticeships often have had a certain amount, and sometimes a substantial amount, of instruction in the trade area they wish to enter. It is important for the industrial personnel manager to know how much of this instruction has been retained by the applicant, because the boy who has profited most by the instruction he has received in his chosen area is most likely to progress with further industrial training toward becoming a skilled tradesman. Although school grades and recommendations of former teachers furnish some indication of progress already made by an applicant, it is frequently desirable to supplement these sources of information with a well-standardized test covering the area in question. The necessity for using a test of this type is illustrated in Figures 5 and 6 reproduced in Chapter 1 on pages 10 and 11. The first illustration shows that among 112 applicants for the job of machine shop apprentice a considerable number made a score of less than 40 (number of items correct) and a scattered few made a score of less than 20 on the machine shop achievement test. Although most of these applicants had had at least one semester of machine shop instruction, and many had had much more than this amount of training, 25 per cent of them did so poorly on the machine shop achievement test that they were below the lowest 10 per cent of students in vocational classes. It was clear, therefore, that among these applicants were a considerable number who would be very unlikely, even with prolonged training, to become expert machinists. The company concerned set 90 items correct as the critical score for hiring for this job and was able, by this process, to select for apprenticeships boys

who made very rapid strides in developing the necessary skill to become expert machinists.

A similar situation was found in the selection of boys for apprenticeships as electricians. Here again a considerable number of the applicants made scores on the Test of Technical Information in Electricity considerably below what would be considered a poor score even among first semester vocational students of this subject. By adopting a critical score for employment that eliminated these unqualified applicants, the apprentices selected constituted a homogeneous group of well-qualified boys who were able to advance rapidly under the systematic apprentice training offered by the industry.

Transfer of employees

The problem of transfer is one that continually confronts every employment manager. Persons hired in one capacity often wish to be changed to a job that offers, or is believed to offer, greater opportunity for advancement. Every personnel manager is not only willing but anxious to transfer employees wherever possible to jobs where they will have a greater opportunity for development. It often happens, however, that an employee wishes to be transferred to a job that he not only cannot perform satisfactorily at the time of transfer but in which, because of lack of aptitude or capacity, he is never likely to be successful. An example in point is the case of tradesmen's helpers. Quite often an employee after working for a period of years as a tradesman's helper feels that he has acquired the necessary information and skill to assume the responsibilities of the tradesman's job. Often such a helper is able to talk quite glibly in terms common to the trade; and this verbal skill, together with a certain familiarity with a few elementary principles of the trade, may convince the employment manager that the employee seeking transfer is now ready for the promotion. In such cases the use of a standardized achievement test furnishes a highly satisfactory means of determining whether or not the employee is really eligible for

the transfer desired. If his score is well below that which is characteristic of apprentices in this trade, the employment manager not only is justified in refusing to make the transfer but, in explaining the reason for his refusal to the employee, he is supplied with the necessary objective information so that his refusal cannot be attributed to prejudice or lack of understanding. Not long ago one employment manager made the comment that the most difficult part of his job was in saying "no" to an applicant seeking employment or an employee seeking transfer. This employment man felt that although he was usually right in his judgment, it was very difficult for him to explain the reasons for his decision to the employee or applicant. It is interesting to note that the judicious use of test results solved his problem. Whenever a request is refused because an employee or applicant has a very low score on an objective test that measures his capacity for the job he is seeking, the person concerned is much less likely to feel that he has been discriminated against.

The use of tests in connection with promotion and transfer supplements rather than replaces such factors as seniority, being "in line" for the job, and the needs of the business.

Discovery of areas needing training

Many industries are devoting more and more time to the systematic training of both new and old employees. The need for such training is a natural result of the continuous technological changes occurring in modern industry. No matter how well qualified an employee may be today, technological change in methods or processes may require that he be completely retrained tomorrow. The systematic use of technical information tests among present employees furnishes a convenient means of determining those areas in which training is needed. An example of this use of information tests is shown in Table 16. This table consists of a set of matching items to determine whether the employees are familiar with the color code used in the plant in which they work. Several

other important areas were covered in the test. The content of the training program that followed the administration of this test was based largely upon the results obtained.

TABLE 16

PART OF AN ACHIEVEMENT TEST COVERING INFORMATION THAT SHOULD BE KNOWN TO EMPLOYEES SEEKING TRANSFER OR PROMOTION

Instructions: On the right is a list of colors. On the left is a list of the materials carried in pipes in this mill. You are to show how well you know the color code by matching each color with the figure or figures which you find before the appropriate materials. Mark your choice in the parentheses at the extreme right. The first one is correctly marked to show you how it should be done. There will be some colors unused. Some others will be used twice.

Materials	*Colors*		
1. Stabilized Gas.................	Aluminum	(6)	( )
2. Steam.........................	Black	( )	( )
3. Hot Water.....................	Brown	( )	( )
4. Cold Water....................	Bright Red	( )	( )
5. Coke Oven Gas.................	Ceiling Blue	( )	( )
6. De-oxidized Gas...............	Dark Purple	( )	( )
7. Compressed Air................	Dark Red	( )	( )
8. Natural Gas...................	Gray	( )	( )
9. Fuel Oil......................	Green	( )	( )
10. Sulphuric Acid...............	Blue	( )	( )
11. Farval Grease................	Pink	( )	( )
12. Bowser Oil...................	Olive Green	( )	( )
13. Gasoline.....................	Orange	( )	( )
14. Kemp Lines...................	Lavender	( )	( )
15. Water Fire Lines.............	White	( )	( )
16. Foamite......................	Yellow	( )	( )
17. Hydraulic Lines..............			

Measurement of vocational achievement

Every vocational teacher expects to place his students in industrial jobs. The success of these students in such jobs, however, depends largely upon the adequacy of training they have received while students. Teachers in the field of general education have long made use of standardized achievement tests to determine student achievement in the several school subjects and to compare the achievement of students in various school systems, under different types of instruction, and in different geographical localities. Administrators and teachers of vocational subjects are now beginning to make similar use of objective achievement tests, and there is every

reason to believe that the judicious use of such tests in this area will be fully as valuable as in the field of general education. The tests of machine shop, electricity, and industrial mathematics discussed in the preceding section are ideally suited for this purpose, and achievement tests in other vocational areas are now in the process of construction. Since vocational demands differ from one industry to another and from one industrial center to another, it is often wise to build tests that are "tailor made" for the particular situation. Such tests, when constructed through the joint efforts of vocational teachers, school administrators, vocational co-ordinators, and representatives of the industries concerned, furnish an ideal means of facilitating the co-ordination between industries and schools that every community desires to encourage.

6

Tests of Personality and Interest

Employment managers universally recognize the importance of personality traits in employees whom they hire. Indeed, one of their reasons for sometimes being hesitant in adopting psychological tests is that they often think of tests only in terms of intelligence or dexterity, and these tests do not, of course, take into account the more general personality traits of the applicant. An applicant might be very high in mental ability or in manipulative dexterity and yet have a personality that would not only make him unfitted for the job for which he is applying but would also make him a definitely undesirable individual to employ in any capacity.

The foregoing sections, which have dealt with aptitude and achievement tests without regard to the applicant's general personality, do not imply that the more general personality traits are unimportant. Psychologists are the first to recognize the importance of personality traits in helping an employee adapt himself to any job or to any organization. Psychological tests have emphasized tests of specific aptitude because psychologists recognize the importance of job aptitude, as such, aside from personality characteristics, and because, up to the present time, it has proved possible to develop adequate tests of such aptitudes as finger dexterity or intelligence to a greater extent than it has been possible to develop adequate tests in the complicated field of personality.

Within recent years, however, both employment managers and consulting psychologists have increasingly demanded some reasonably satisfactory and accurate method of deter-

mining certain personality traits of an applicant at the time he applies for employment. It is now recognized not only that an applicant who does not have the aptitude to learn the job will fail, no matter how desirable his personality traits may be, but also that if he does have the aptitude for the job he will probably still fail if his other personality characteristics make it difficult for him to fit into the organization and to work cooperatively with other persons. Although this conclusion is sound common sense, we do not need to rely entirely upon subjective judgment to reach it. Such investigations as those of Hunt[1] and Brewer,[2] which have been carried on with thousands of employees in a variety of industries, show that personality factors, rather than lack of ability on the job, are responsible for a large number of layoffs and failures to be promoted. These studies deal with layoffs due to factors other than failure to produce. The rising voice of labor in determining managerial policies, particularly with respect to hiring and firing, and the growing importance of seniority as a determining factor in layoff, have made it increasingly important to determine, at the time of hiring, whether an applicant has any incipient personality maladjustments that might prevent him from fitting properly into the organization. It is with the hope of developing some sort of test or tests to accomplish this purpose that a number of psychologists have developed scales for the measurement of personality or temperament traits.

Tests of Personality

Probably the first attempt to develop a scale of this type was made by Downey[3] and resulted in the Downey Will-Temperament Scale. This test was based on a number of

---

[1] H. C. Hunt, "Why People Lose Their Jobs or Aren't Promoted," *Personnel Journal*, XIV (1935–1936), p. 227.

[2] J. M. Brewer, "Religion and Vocational Success," *Religious Education*, XXV (1930), pp. 29–41.

[3] J. E. Downey, "The Will-Temperament and Its Testing" (World Book Company, New York, 1923).

assumptions that were later found to be untenable, and for that reason the scale was never very widely used. It is noteworthy, however, in that it served the purpose of directing the attention of psychologists to the need for a measuring instrument in this area. Some time later the Thurstone Personality Schedule[4] was published. This scale gives a single gross score indicating the presence or absence of neurotic tendencies. With this scale, Thurstone was able to show that neurotic tendencies are relatively independent of mental ability, but are related to accomplishment in certain areas, particularly college and university work. An abbreviated form of Thurstone's scale has been published by Willoughby.[5] Following the early work in the field of personality testing, a number of comprehensive scales for the measurement of various aspects of personality have been developed.

The Bernreuter Personality Schedule[6]

This is one of the most widely used personality tests. It consists of 125 questions dealing with a wide variety of topics that indicate personality traits and characteristics. The scale may be scored with each of six different scoring keys, each of which measures a different personality component. The components measured, with the codes used in referring to them, are:

B1–N. A measure of neurotic tendency. Persons scoring high on this scale tend to be emotionally unstable. Those scoring above the 98 percentile would probably benefit from psychiatric or medical advice. Those scoring low tend to be very well balanced emotionally.

B2–S. A measure of self-sufficiency. Persons scoring high on this scale prefer to be alone, rarely ask for sympathy or encourage-

---

[4] L. L. Thurstone and T. G. Thurstone, Personality Schedule (University of Chicago Press, 1929).

[5] R. R. Willoughby, Thurstone Personality Schedule, Clark Revision (Worcester, Mass., 1932).

[6] R. G. Bernreuter, The Personality Inventory and Manual (Stanford University Press, 1935).

ment, and tend to ignore the advice of others. Those scoring low dislike solitude and often seek advice and encouragement.

B3–I. A measure of introversion-extroversion. Persons scoring high on this scale tend to be introverted; that is, they are imaginative and tend to live within themselves. Scores above the 98 percentile bear the same significance as do similar scores on the B1-N scale. Those scoring low are extroverted; that is, they rarely worry, seldom suffer emotional upsets, and rarely substitute day-dreaming for action.

B4–D. A measure of dominance-submission. Persons scoring high on this scale tend to dominate others in face-to-face situations. Those scoring low tend to be submissive.

F1–C. A measure of confidence in oneself. Persons scoring high on this scale tend to be hamperingly self-conscious and to have feelings of inferiority; those scoring above the 98 percentile would probably benefit from psychiatric or medical advice. Those scoring low tend to be wholesomely self-confident and to be very well adjusted to their environment.

F2–S. A measure of sociability. Persons scoring high on this scale tend to be non-social, solitary, or independent. Those scoring low tend to be sociable and gregarious.

Scoring keys for the first four of the components (indicated by code *B*) were developed by Bernreuter[7] in his original work with the scale. The last two (indicated by code *F*) were derived from a factor analysis of the Bernreuter scale reported by Flanagan.[8]

Because of the long time required to score the Bernreuter test with the original scoring keys, a simplified scoring system has been developed and reported by McClelland.[9] The simplified scoring system proposed gives scores that, for components B1-N, B4-D, F1-C, and F2-S, correlate above .95 with scores obtained by the original scoring method. For component B2-S the correlation is .84. A simplified scoring

---

[7] R. G. Bernreuter, "The Theory and Construction of the Personality Inventory," *Journal of Social Psychology*, IV (1933), pp. 387–405.

[8] J. C. Flanagan, *Factor Analysis in the Study of Personality* (Stanford University Press, 1935).

[9] D. C. McClelland, "Simplified Scoring of the Bernreuter Personality Inventory," *Journal of Applied Psychology*, XXVIII (1944), pp. 414–419.

for component B3-I was not developed because the latter correlates so highly with B1-N.[10]

Several investigations conducted in business and industrial situations have been reported that show the effectiveness of the Bernreuter Personality Schedule. Richardson and Hanawalt[11] found that the Bernreuter scales differentiate between "office holders" (men who have had at least two presidencies or important chairmanships since the age of twenty-one in business, professional, civic, religious, fraternal, or social organizations) and men who have not held such offices.

The "office holders" made scores which showed them to be significantly more dominant, more self-confident, less neurotic, and less introverted than those who had held no office. In this same study, Richardson and Hanawalt reported significant differences between the Bernreuter scores of supervisors (men who had fifteen or more persons under their direction or supervision) and those of a contrasting non-supervisory group. The supervisors made scores showing them to be less neurotic, less introverted, more dominant, and more self-confident than the non-supervisory group.

The Humm-Wadsworth Temperament Scale[12]

This scale has been developed by two men whose work has been primarily in the field of industrial personnel. The scale consists of 318 questions which the person tested answers by checking *yes* or *no*. From the answers to these questions it is possible, by differential scoring, to obtain separate scores for seven aspects of temperament. These aspects are as follows:

---

[10] R. G. Bernreuter, *Manual for the Personality Inventory*, (Stanford University Press, 1935).

[11] H. M. Richardson and N. G. Hanawalt, "Leadership as Related to the Bernreuter Personality Measures: III. Leadership Among Adult Men in Vocational and Social Activities," *Journal of Applied Psychology*, XXVIII (1944), pp. 308–317.

[12] D. G. Humm and G. W. Wadsworth, The Humm-Wadsworth Temperament Scale, Test Booklet and Manual, second 1940 revision (Doncaster G. Humm Personnel Service, Los Angeles, California).

1. **The Normal Component.** This is primarily a control mechanism providing rational balance and temperamental equilibrium. It underlies the conservatism, toleration, and conformity to socially acceptable behavior observed in the well-adjusted person.

2. **The Hysteroid Component.** An individual with an excess of the hysteroid component possesses a character defect with ethically inferior motivation manifested by stealing, lying, cheating, and similar antisocial behavior.

3. **The Manic Cycloid Component.** This is characterized by emotionality, fluctuation in activities, and interferences with voluntary attention, some degree of elation, pressure of activity, and distractibility together with such manifestations of excitement as jests, pranks, enthusiasms, impatience, and so forth.

4. **The Depressive Cycloid Component.** This is manifested by some degree of sadness, lessened activity, and associated characteristics such as worry, timidity, and feeling of malaise.

5. **The Autistic Schizoid Component.** This is characterized by heightened imagination, leading to a tendency toward a daydream life concerning which the subject is sensitive.

6. **The Paranoid Schizoid Component.** This includes stubborn adherence to fixed ideas, suspicion, and contempt for the opinion of others, with behavior fitting these traits.

7. **The Epileptoid Component.** This is characterized by inspirations to achievement that are meticulously developed and pushed through to completion.

Everyone who studies the above list of components with their brief descriptions will recognize certain persons of his acquaintance who, from long observation, are clearly known to possess an excessive amount of one or another or of some combination of these characteristics. One does not need the Humm-Wadsworth Scale, or perhaps even training in psychology, to identify the Hysteroid or the Cycloid if he is constantly thrown into contact with an individual of this type. Indeed, one thrown into contact with such a person is more likely to recognize that something is wrong than is the afflicted individual himself. The mentally ill or "near ill" often feel that they are quite normal but that everybody else is wrong. It is, however, one thing to recognize the presence of such personality characteristics in a person with whom one

is thrown into constant contact, and it is quite another thing to recognize them in an applicant whom one is interviewing for the first time. It is claimed that the Humm-Wadsworth Temperament Scale will accomplish this result.

It is, however, expecting a good deal of any scale of this type to identify such personality traits in persons who are "after" jobs and who can therefore be expected to answer the questions in whatever way they feel the questions should be answered in order to be sure they get the jobs. We have already mentioned that the questions are of a type that can be answered either *yes* or *no*, and that an applicant for a job may not be too honest if he feels that honest answers will decrease the chances of his being employed. The extent to which the

TABLE 17

MEAN SCORES ON HUMM-WADSWORTH TEMPERAMENT SCALE OBTAINED IN A CLINICAL AND AN EMPLOYMENT SITUATION BY 65 COLLEGE STUDENTS*

Component	*Clinical Mean*	*Employment Mean*	*Shift from Clinical to Employment*
"Normal"	981	1023	+42
Hysteroid	1023	980	−63
Manic	1035	937	−98
Depressive	1061	913	−148
Autistic	1024	938	−86
Paranoid	970	955	−15
Epileptoid	983	1002	+19

* All scores were computed by the log method, with correction for no-count, as described in the second 1940 revision of the manual of directions.

scores on the seven components may shift when a person is changed from a frank or clinical situation to a job-application situation is revealed in a study conducted with 65 college students.[13] Each student was given the scale twice: first, with instructions to be as frank as possible, and second, to assume that he was in an employment office after a job and had been asked to take the test as a part of the employment procedure. Table 17 shows the mean scores for the seven components obtained under these two conditions.

[13] This study was conducted by W. J. Giese and F. C. Christy at Purdue University.

It is apparent from Table 17 that the employment situation, when compared with the clinical, shows a higher average value for the normal component and lower values for all except the epileptoid of the remaining six components. In other words, the students were able, by assuming an attitude of "applying for a job," to change their test profiles toward more of the normal and less of the undesirable traits. All differences shown in Table 17 are significant from a statistical viewpoint, and only those scores were used in the computation that fell within the no-count limits, for both the clinical and employment situations, within which the manual of instructions states that the scale should be "accepted as probably valid."

Even with the shift in means from one situation to the other, it would still be possible to infer one's score in one situation from a knowledge of his score in the other situation if scores in the two situations were highly correlated. The correlations were computed and are shown in Table 18.

TABLE 18

CORRELATIONS BETWEEN CLINICAL AND EMPLOYMENT SITUATIONS ON THE HUMM-WADSWORTH TEMPERAMENT SCALE FOR 65 COLLEGE STUDENTS

Component	*Correlation*
"Normal"	−.03
Hysteroid	+.42
Manic	+.09
Depressive	−.10
Autistic	+.11
Paranoid	+.61
Epileptoid	+.23

The only correlation in Table 18 that is large enough to be significant is the one for the Paranoid Component.

We are thus forced to conclude that the Humm-Wadsworth Temperament Scale will not give a completely invariable picture of a person under any and all conditions of testing or point of view of the applicant. This criticism should not, of course, be made specifically of the Humm-Wadsworth test, but of personality tests in general. It is probably true that *any* test which allows the subject to answer according to his

judgment or feelings is open to the same possibility of falsification. This fact further emphasizes the importance of "testing the test," as described in Chapter 3, before attempting to make practical use of it in any specific situation.

The Guilford Series of Personality Tests

Guilford and his collaborators[14] have conducted an extensive series of statistical investigations that have identified a number of personality factors. These investigators have also devised scales to measure several of these factors identified. The Guilford series of three personality scales follows:

An Inventory of Factors S T D C R[15]

This scale measures the factors:

S—Social introversion-extraversion.—Shyness, seclusiveness, tendency to withdraw from social contacts, versus sociability, tendency to seek social contacts and to enjoy the company of others.

T—Thinking introversion-extraversion.—An inclination to meditative or reflective thinking, philosophizing, analysis of one's self and others, versus an extravertive orientation of thinking.

D—Depression.—Habitually gloomy, pessimistic mood, with feelings of guilt and unworthiness, versus cheerfulness and optimism.

D—Cycloid disposition.—Strong emotional fluctuations, tendencies toward flightiness and emotional instability, versus uniformity and stability of moods, evenness of disposition.

R—Rhathymia.—A happy-go-lucky, carefree disposition, liveliness, impulsiveness, versus an inhibited, over-controlled, conscientious, serious-minded disposition.

Guilford-Martin Inventory of Factors G A M I N[16]

This scale measures the factors:

G—General pressure for overt activity.

---

[14] J. P. and R. B. Guilford, "Personality Factors S, E, and M, and their Measurement," *Journal of Psychology*, II (1936), pp. 107–127; "Personality Factors D, R, T, and A," *Journal of Abnormal and Social Psychology*, XXXIV (1939), pp. 21–36; and "Personality Factors N and GD," *ibid.*, XXXIV (1939), pp. 239–248.

[15] Distributed by Sheridan Supply Co., Beverly Hills, Calif.

[16] *Ibid.*

A—Ascendancy in social situations as opposed to submissiveness; leadership qualities.

M—Masculinity of attitudes and interests as opposed to femininity.

I—Lack of inferiority feelings; self-confidence.

N—Lack of nervous tenseness and irritability.

The Guilford-Martin Personnel Inventory I[17]

This scale measures the factors:

O—Objectivity (as opposed to personal reference or a tendency to take things personally)

Ag—Agreeableness (as opposed to belligerence or a dominating disposition and an overreadiness to fight over trifles)

Co—Cooperativeness (as opposed to faultfinding or overcriticalness of people and things)

The three components of this inventory all measure certain aspects of the paranoid trait of temperament. Since an excessive amount of this trait has long been considered to underlie many of the difficulties encountered by some people in situations requiring contact with others, this scale would seem to have ideal possibilities for use in industrial situations. An indication of its validity in an industrial situation is given by the results obtained in a Southern California industry.[18]

The test was given to 51 employees, among whom were 22 who were regarded by management as trouble makers or malcontents. On the basis of test scores, and without knowledge of the ratings, the employees were divided into a satisfactory and an unsatisfactory group. Since 22 of the original 51, or 43 per cent, had been rated as troublemakers, it would be expected on a chance basis that *any* testing technique would identify unsatisfactory employees in 43 per cent of the trials. However, only 34 per cent of the high scoring group had received unsatisfactory ratings, as contrasted with the 73 per cent of the low scoring group who had

---

[17] *Ibid.*

[18] The Guilford-Martin Personnel Inventory, Manual of Directions and Norms (Sheridan Supply Co., Beverly Hills, Calif.).

received them. This discrepancy indicates that the identification of problem employees can be achieved by this scale with considerably greater accuracy than could be expected of identification on a chance basis.

The Minnesota Multiphasic Personality Inventory[19]

This inventory consists of 550 statements that the person being tested sorts into three piles—those he regards as *true*, those he regards as *false*, and those on which he cannot make a true or false judgment. After the cards have been sorted, the scoring procedure yields eight different personality component scores, viz:

1. **Hypochondriasis Scale.** Measures undue worry about one's health.
2. **Depression Scale.** Measures feeling of unworthiness, uselessness, and undue worry.
3. **Hysteria Scale.** Measures the degree to which subject is like patients who have general systemic complaints and/or specific complaints such as paralyses, contractures, and gastric, intestinal, and/or cardiac symptoms.
4. **Psychopathic Deviate Scale.** Measures similarity of subject to patients whose main difficulty lies in their absence of deep emotional response, inability to profit from experience, and disregard of social moves.
5. **Interest Scale.** Measures tendency toward masculinity or feminity of one's interest pattern.
6. **Paranoia Scale.** Measures stubborn adherence to fixed ideas, suspicion, and contempt for the opinion of others.

---

[19] This test and the accompanying manual are distributed by the University of Minnesota Press, Minneapolis, Minn. The research on which the scales are based is reported in the following articles:

S. R. Hathaway and J. C. McKinley, "A Multiphasic Personality Schedule (Minnesota): I. Construction of the Schedule," *Journal of Psychology*, X (1940), pp. 249–254; "II. A Differential Study of Hypochondriasis," *ibid.*, X (1940), pp. 255–268; "III. The Measurement of Symptomatic Depression," *ibid.*, XIV (1942), pp. 73–84; "IV. Psychasthenia," *Journal of Applied Psychology*, XXVI (1942), pp. 614–624.

J. C. McKinley and S. R. Hathaway, "The Minnesota Multiphasic Personality Inventory: V. Hysteria, Hypomania, and Psychopathic Deviate," *ibid.*, XXVIII (1944), pp. 153–184.

7. Psychasthenia Scale. Measures similarity of the subject to psychiatric patients who are troubled with phobias (unreasonable fears) or compulsions.

8. Schizophrenia Scale. Measures similarity of subject to psychiatric patients who are characterized by bizarre and unusual thoughts and behavior.

In addition to the eight personality components measured, the Minnesota Multiphasic Personality Inventory also provides for certain checks on the accuracy and honesty of the subject's responses. A further expansion of the scoring possibilities of the scale has been developed by Drake[20] in the form of a key to obtain scores on social introversion-extroversion. Simplified scoring of the scale for group administration, to be used when electric, punched-card tabulating equipment is available, has been described by Burton and Bright.[21]

An indication of the validity of the Minnesota Inventory comes from the work of Schmidt,[22] who reports that the Inventory "distinguished . . . much statistical significance between normal soldiers and those diagnosed as constitutional psychopaths; mild or severe neurosis; and psychosis."

Cautions in the use of personality tests

There is a great temptation to use one or more of the personality tests without taking necessary precautions in their interpretation. The most common type of mistake is to reason as follows:

1. One personality component measured by a given test is called *neurotic tendency*. This test *must* measure *neurotic tendency* because that title is used by the authors of the test. (The reader may substitute any presumably undesirable personality trait he wishes for the

---

[20] L. E. Drake, "A Social I. E. Scale for the Minnesota Personality Inventory," *Journal of Applied Psychology*, XXX (1946), pp. 51–54.

[21] A. Burton and C. J. Bright, "Minnesota Multiphasic Personality Inventory for Group Administration," *Journal of Consulting Psychology*, X (1946), pp. 99–103.

[22] H. O. Schmidt, "Test Profiles as a Diagnostic Aid: The Minnesota Multiphasic Inventory," *Journal of Applied Psychology*, XXIX (1945), pp. 115–131.

phrase *neurotic tendency.* The result of the "reasoning" will be the same.)

2. We do not want in our plant people who have a high degree of *neurotic tendency.*

3. Therefore we do not want in our plant anyone who scores beyond a certain point on this test.

Anyone familiar with the statistical procedures that should be followed in developing and standardizing any test for use in a given situation will recognize a fallacy in every one of the three steps given above. First, a test does not necessarily measure a given trait merely because the name of that trait is used in the title or description of the test. Second, one cannot assume that persons with *any* given trait, *as revealed by a test score on a test,* are by virtue of that fact alone necessarily undesirable employees. It follows that the third statement, based upon the accepted truth of the first and second, is hardly justified without more specific evidence on the applicability of the test to the situation in question.

The validity of a personality test must be determined exactly as the validity of any other test must be determined, namely, by the methods of "testing the test" described in Chapter 3. Unless the test is proved to measure something that is important for job success, it should not be used, regardless of what it is supposed to measure, and regardless of the words or phrases used in its title or description.

Entirely aside from the use of personality scales as a device for selecting employees, the possibility exists of using them as a means of helping present employees overcome possible weaknesses that may exist in their temperaments. A number of industrial men have reported the successful use of personality scales for this purpose. The procedure consists in having the test taken by those employees or supervisors who are having some sort of trouble in adjusting their personalities to the demands of the situation. Very often it is possible to awaken among employees an interest in self-improvement that will inspire them to take the personality test as an experi-

mental venture. Under such circumstances, they often give rather frank and honest answers to the questions. The resultant profiles bring into sharp relief any personality deviations that may be present. With the profile before him, a skillful counselor can often bring the person tested to realize that the picture presented by the test is fairly accurate. When this much has been accomplished; when the employee has been given an insight into his weaknesses as well as his strengths, he is much more able to do something about himself than if he goes about from day to day without realizing that certain aspects of his behavior may be "rubbing people the wrong way."

For such a purpose, it is not necessary that the test give a completely accurate representation of an individual's personality. If it brings into relief even partially those aspects of his personality which most need attention, and if a skillful counselor can bring about an insight into this condition without upsetting the employee's general emotional balance, a very real step in the direction of self-improvement has been made. Many industries have found personality tests to be a real help in connection with such counseling procedures within their plant.

The Measurement of Interest

Everyone realizes that an individual's interests direct his activities. The student who is interested in engineering will study long hours, apply himself with diligence, and achieve a mastery of engineering subjects that is limited only by his capacity to master those subjects. A student who lacks this interest, even though he may have the same or even greater capacity, will find excuses for not studying his calculus, will go to the movies or to a dance when he should be writing up laboratory reports, and will in other ways avoid those hours of prolonged study without which successful accomplishment cannot be attained.

Vocational Guidance counselors have long recognized that

it is just as important to be sure a boy is studying for a vocation in which he has a real and driving interest as to be sure that he has the capacity to achieve success in that field. Without an interest, no amount of capacity will be sufficient; just as without the capacity, no amount of interest will be sufficient.

Up to the present time, very little has been done in the industrial placement of employees by way of ascertaining their interests and placing them upon jobs for which they not only have the necessary capacity but in which they also are definitely interested.

The accurate measurement of the true interests of persons applying for jobs is subject to the same difficulties as the measurement of their personalities. If one is very anxious to obtain a certain job (for monetary or other considerations), it is not likely that he will reveal his true interest if that interest happens to be in some direction that is not related to the job. The industrial application of interest-measuring techniques is therefore limited largely to situations in which the person tested is not to be selected or rejected for employment as a result of the test, but rather is to be hired anyway and will be placed or later transferred in accordance with his basic interests. Several interest tests have been developed and published.

Strong Vocational Interest Blank[23]

This test has been developed primarily for use in vocational guidance counseling. It determines whether the subject's pattern of interests agrees with the interest pattern of men in each of a large number of professions and occupations. Other things being equal, a person choosing a profession or occupation is more likely to be happy and successful if his basic interests are similar to the interests of men actually in this particular field.

[23] E. K. Strong, Jr., "Vocational Interest Test," *Educational Record*, VIII (1927), pp. 107–121. This test is distributed by the Stanford University Press, Stanford, Calif.

Different forms of this test and different scoring keys are available, so that it may be used with men or women, with students or adults.

The Strong Vocational Interest Blank measures as many patterns of interest as vocational patterns for which it is scored. This procedure is advisable in vocational guidance work, and the test is admirably suited for that purpose. Recent work, however, has indicated that interests may be grouped into patterns, the number of which are considerably smaller than the possible number of vocations from which a choice must be made. Thurstone[24] reports a multiple factor study of vocational interests that shows by an analysis of Strong Blank results that interests may be basically divided into four general fields: science, language, people, and business. It would seem to be important for the personnel manager to consider these interests as well as abilities, particularly in the case of college graduates or others who are expected to show considerable development over a period of years. It is probably unwise to place in the scientific development department young college men whose basic interests are in people and business; and it would seem to be equally unwise to put into the sales department those whose basic interests are scientific. Fortunately, it is often possible to obtain a fairly accurate picture of an applicant's or employee's interests by means of an interview in which a number of questions are asked about such topics as present and past activities, hobbies, and how vacations are spent.

Kuder Preference Record[25]

This is an interest test that is scored in terms of basic interest groups, as described above, rather than in terms of

[24] L. L. Thurstone, "A Multiple Factor Study of Vocational Interests," *Personnel Journal*, X (1931), pp. 198–205.

[25] G. F. Kuder, "The Stability of Preference Items," *Journal of Social Psychology*, XIX (1939), pp. 47–50. This test is distributed by Science Research Associates, 228 South Wabash Ave., Chicago, Ill.

interest patterns of various specific vocations. The interest groups covered by the Kuder Preference Record are:

Mechanical
Computational
Scientific
Persuasive
Artistic
Literary
Musical
Social Service
Clerical

Separate scores are readily obtained for each of the listed interest groups. This type of scoring makes the test particularly easy to try out on a "test the test" basis (see page 52) because only the nine basic scores will be considered, and these yield an over-all pattern of interests. The manual with this test, however, provides interpretive data for evaluating various interest patterns in terms of their agreement with those of men in various occupations.

Cardall Primary Business Interests Test[26]

This test is designed to measure an individual's preferences for the specific job activities that characterize beginning business jobs. These immediate and specific preferences point to the initial job, determine the individual's interest or boredom in his first activities, and often determine to a considerable extent his progress in his work. The five business fields in which this blank measures interest are: accounting, collecting and adjusting, sales office work, sales store work, and stenographer-filing.

Several investigations have indicated that the interest questionnaires, when used in this manner, are of value in

---

[26] A. J. Cardall, "A Test for Primary Business Interests Based on a Functional Occupational Classification," *Educational and Psychological Measurements*, II (1942), pp. 113–138. This test is distributed by Science Research Associates, 228 S. Wabash Ave., Chicago, Ill.

industry. Shartle[27] made a study of two groups of foremen, one consisting of highly successful supervisors and the other made up of foremen approximately equal in job skill but less able than the first group in handling the personal aspects of their supervisory jobs. In Shartle's investigation the Strong Vocational Interest Blank[28] was given to all the supervisors, and this was followed by a detailed clinical interview. The results showed that the less successful foremen were characterized by more interest in withdrawing from others, by more indifference to the actions of others, and by more antagonistic reactions toward others. The study suggests that interest in dealing with others is one of the prime requisites of the successful supervisor. Laycock and Hutcheon[29] have reported a study in which a battery of tests containing, among other things, a measure of interest in physical science gave a correlation of .66 with success in engineering courses.

While it is not always possible for an employment manager to place employees according to their interests, it is perhaps possible for him to do so more frequently than is often realized. In the shifting and transfer that continually occur in any large business or industrial plant, it is good business as well as good industrial relations to consider whenever possible the basic interests and desires of employees who are being shifted. Adequate consideration of these interests goes deeper than basing transfers upon mere statements of preference. It is entirely possible that in many cases the employee himself does not know what type of work he would be most interested in because he is not sufficiently familiar with various types of work. Interest tests measure an individual's basic interests by sampling a number of activities that represent different kinds of jobs. Such a test gives a reasonably accurate

[27] C. L. Shartle, "A Clinical Approach to Foremanship," *Personnel Journal*, XIII (1934), pp. 135–139.

[28] Strong, *op. cit.*

[29] S. R. Laycock and N. B. Hutcheon, "A Preliminary Investigation into the Problem of Measuring Engineering Aptitude," *Journal of Educational Psychology*, XXV (1939), pp. 280–289.

indication of the individual's basic interest in various jobs even though he is not familiar with the details of the job in question.

Much remains to be done in the fields of personality and interest tests. A serious obstacle in both these fields is the employee who does not want to be measured, or who wants to be falsely measured, and who may not, therefore, give useful test results. Certainly this difficulty is a much greater obstacle to personality and interest measurement than to the measurement of intelligence, dexterity, or trade achievement. An employee whose mental ability is such that he is able to answer correctly only ten problem questions out of fifty asked is utterly unable to answer twenty or thirty of the questions in order to give the false impression that he has more mental ability than he really has. The test itself sets the limit of his performance; and while he may have *more* ability than is shown on the test (because of being ill or otherwise indisposed when the test was taken), we can safely assume that he does not have *less* ability. But in the measurement of both personality and interest it is necessary to be sure, before the test is administered, that the employee himself has been convinced that sincerity and truthfulness in answering the questions will operate to his own eventual advantage in helping the employment manager place him where he is most likely to achieve success. Unless the situation is such that the employee can be so convinced, it is seldom wise to place too much confidence in the results of the personality and interest tests that are now available.

7

Visual Skills and Vision Tests

PRACTICALLY every industrial job requires some degree of vision, and many jobs require a high degree of skill in some particular visual function. The inspector of small parts for "appearance" must have keen vision at close distances. The operator of certain knitting and other textile machines must not only have keen vision at close distances but also must be able to maintain such vision for long periods of time with only occasional interruption. The truck driver, crane operator, and signal man must have keen vision for greater distances and good perception of space relationships. Color discrimination is of importance to an employee wiring a radio (because he must discriminate between wires of different color), to the operator of a color-printing press, to the pipefitter tracing a color code through the plumbing system of a plant, and to the operator of mobile equipment who must depend upon colored signals to determine whether roads are open or pathways clear. Various measurable visual characteristics have been found to be related to successful performance on certain jobs—even on jobs in which these visual factors could not be inferred from ordinary job analysis procedures. The pattern of visual requirements differs markedly for different jobs, and these requirements cannot be discovered except by a study of the people on the job, as explained in Chapters 1 and 3.

Management in productive industries has long recognized the importance of vision in employees.[1] Some form of vision

[1] Extensive recent publications in this field are listed in bibliographies on industrial vision:

H. S. Kuhn, "Articles Bearing on Industrial Eye Problems," *Transactions*

test, administered either in the employment office or as a part of the medical examination, is perhaps more common in industry than is any other form of employee test. Testing visual skills has always been a legitimate psychological function, just like testing carried on in other areas of skill or aptitude. Visual skills, however, are different in several respects from other human capacities that are measured by psychological tests. Visual skills are not innate and permanent; they are to a large degree acquired and can be, and often are, modified. They change universally with age. Professional eye care, based on a clinical or diagnostic examination, can often give to an employee the visual skills he needs for his job or restore the skills he has lost with increasing age.

Measurement of Vision

Measurement of visual skills of industrial employees and applicants may be undertaken for any of several purposes. It is important for the industrial psychologist to know the various purposes and points of view that are represented in existing visual programs in industry so that he may relate his own work to the activities already under way. More critical and comprehensive tests of vision are necessary, designed rather for economically valid selection, placement, and improvement of employees than for some of the other, simpler objectives of vision testing that have been attained in industry heretofore.

Visual skills

The investigations reported in this chapter have clearly indicated the need for adequate visual classification and placement tests *specifically adapted to industrial purposes and conditions.* These investigations in many types of industry have indicated certain aspects of visual performance as being

of the American Academy of Ophthalmology and Otolaryngology, 1946, pp.175–178.

N. F. Stump, "A New Concept of Ophthalmic Service," *Bausch and Lomb Magazine*, XXII (1946), No. 1, pp. 4–20.

of most general importance for classifying and placing employees according to differences in visual characteristics.

1. *Keenness of vision* (*visual acuity*) *at appropriate distances*—usually tested at twenty feet and thirteen or sixteen inches. This visual function is the ability to discriminate black and white detail, measured in terms of the minimum separable areas that can be distinguished. For industrial placement such a test should be equally valid for illiterate and literate subjects and should avoid the complications introduced by a factor involving discrimination and recognition of different shapes, such as letters. The scale should make it possible to classify acuity scores adequately throughout the entire range of performance. Separate scores should be recorded for both eyes together and each eye alone. Since acuity is modified by brightness, glare, and other external conditions, the acuity test must be given under controlled and standard conditions.

2. *Discrimination of differences in distance* (*depth perception, or stereopsis*). This function is an important phase of correct perception of spatial relationships. Of several cues for judging relative distances of objects, the most important for normal two-eyed persons, and the one that can be controlled and measured most reliably, depends on the slight difference in the position of the two eyes. The two eyes perform a geometric triangulation upon a distant object, and the distance of that object is perceived through an integration of the minute differences in appearance of the object to the two eyes. Other cues for perceiving distance in the third dimension may augment but cannot adequately substitute for this cue from two-eye functioning. Stereopsis is measurable quantitatively like most other human functions and should be so measured for employee placement.

3. *Discrimination of differences in color*. Accurate color discrimination is important on some jobs, and it may also reflect certain aspects of health that are important for adequate performance on many jobs. The particular colors to be differentiated should include as many combinations as

there are factors in color sensitivity. The most important combinations are those that represent common colors in signal lights. In order to read such signals correctly, it is important not only that an employee should be able to differentiate between them but also that he should be able to identify and interpret correctly the meaning of each color. Any test of color vision can and should be scaled for quantitative measurement and classification of employees.

4. *Postural characteristics of the eyes* (*phorias*) *at appropriate distances*—usually twenty feet and thirteen or sixteen inches. Under certain testing conditions, which eliminate the necessity for the eyes to converge on a single point, the eyes assume a posture that may converge or diverge from that required in normal seeing at the test distance. Such postures (called "phorias" in clinical terminology) are measured in terms of angular deviation from the posture normally required for that distance. The deviation may be lateral or vertical and is measured separately in each direction. The explanation of this phenomenon has not been determined finally; nevertheless such characteristics should be measured because they are related to performance on most industrial jobs. Such measurement must be done with adequate control and standardization of the several factors that may modify the measurement, such as the distance and focus requirement of the test.

These are not, of course, the only visual functions that are of importance in all industries, but they comprise a reasonable minimum for a program of testing applicants and employees. These tests have been adapted satisfactorily for simple and rapid use in industry, and they are the ones most widely recommended by leaders in industrial eye care[2] for use in industry.

---

[2] Hedwig S. Kuhn, *A New Concept of Visual Performance in Industry.* Publication No. 340 (National Society for the Prevention of Blindness, 1940).

A. C. Snell, "The Field of Industrial Ophthalmology," American Medical Association. Section of Industrial Health (Chicago, Jan. 13, 1942).

C. F. Shepard, "Visual Skills," *Optometric Weekly*, XXXIV (1944), pp. 1465–1466.

The Bausch and Lomb Visual Classification and Placement Tests for Industry[3] are the first battery of vision tests to be constructed on the basis of specifications derived from extensive investigations among industrial employees in industrial situations. These tests cover the visual functions described above and, for maximum speed and convenience in testing, are incorporated in a single instrument, shown in Figure 49.

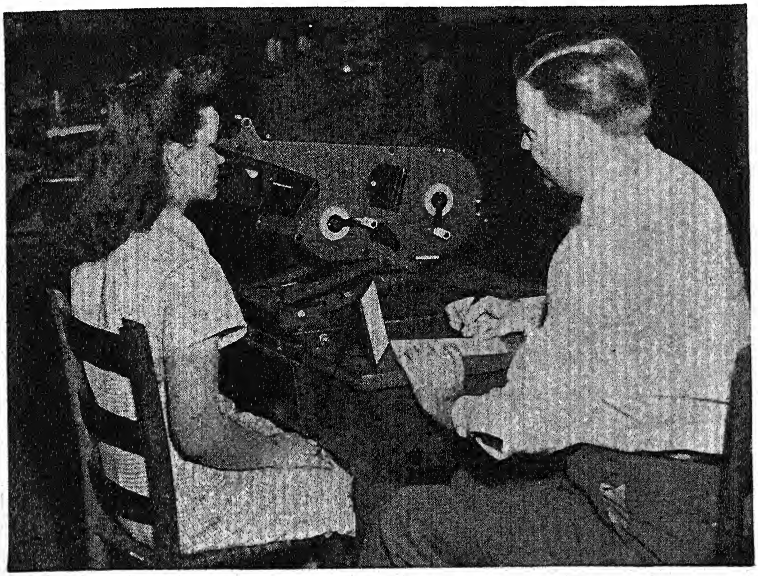


FIG. 49—The Ortho-Rater in use in a hosiery mill. (Courtesy of Haynes Hosiery Mills Company, Winston-Salem, N. C.)

This instrument, called the Ortho-Rater, is a precision stereoscope of relatively long focal length that permits adequate and separate control of test stimuli for each eye. Tests are given at optical equivalents of twenty-six feet and thirteen inches. Stereoscopic methods of vision testing have been used since the late nineteenth century and were early described by Wells.[4] The statistical data upon which these

[3] The Bausch and Lomb Optical Company, Rochester, N. Y.

[4] David Wells, *The Stereoscope in Ophthalmology*, second edition (Globe Optical Company, 1918).

tests are based are being published separately in technical and professional journals.[5] A number of validating studies will be cited in this chapter.

Individual differences in vision

Employees and applicants differ just as markedly in visual characteristics, when these are adequately measured,

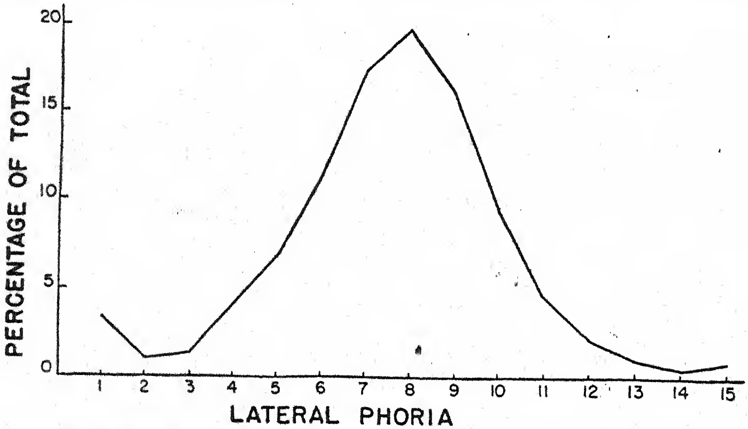


FIG. 50—Distribution of Ortho-Rater scores on far lateral phoria for 3,313 employees.

as in any other, and these differences correlate in many respects with differences in job performance. Differences in near visual acuity have already been shown in Figure 10. Figure 50 shows differences in far lateral phoria for 3,313 employees in a plant manufacturing accounting machines. Figure 51 shows differences in near vertical phoria for 2,394

[5] W. J. Giese, "The Interrelationship of Visual Acuity at Different Distances," *Journal of Applied Psychology*, XXX (1946), pp. 91–106.

F. W. Jobe, "Instrumentation for the Bausch and Lomb Industrial Vision Service," *Bausch and Lomb Magazine*, XX (1944), No. 2, pp. 3–5, 19–21.

Joseph Tiffin and H. S. Kuhn, "Color Discrimination in Industry," *Archives of Ophthalmology*, XXVIII (1942), pp. 851–859.

Joseph Tiffin and S. E. Wirt, "Near *vs.* Distance Visual Acuity in Relation to Success on Close Industrial Jobs," *Supplement to Transactions of the American Academy of Ophthalmology and Otolaryngology*, June, 1944, pp. 9–16.

S. E. Wirt, "The Validity of Lateral Phoria Measurements in the Ortho-Rater," *Journal of Applied Psychology*, XXVII (1943), pp. 217–232.

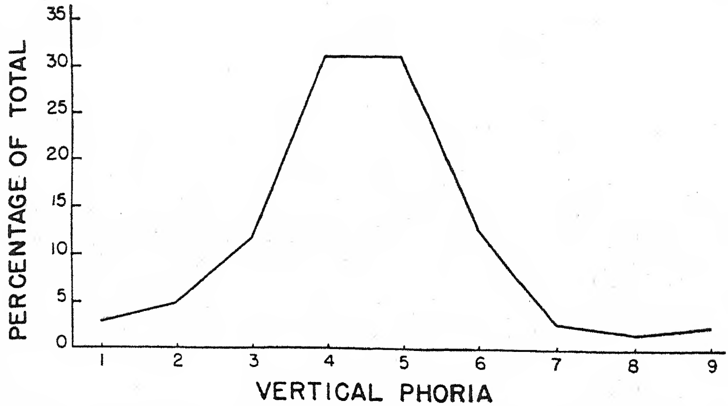


FIG. 51—Distribution of Ortho-Rater scores on near vertical phoria for 2,394 employees.

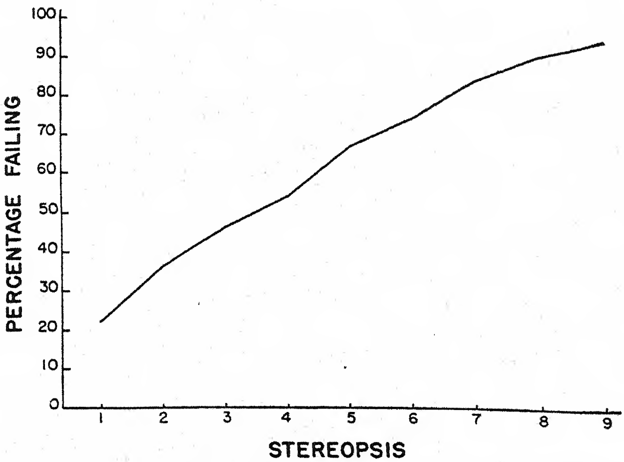


FIG. 52—Percentages of 1,895 employees failing the Ortho-Rater depth perception test at different levels of difficulty.

men employed in an optical factory. These three curves, based on Ortho-Rater tests, show the typical bell-shaped distributions characteristic of measurements of human traits. By such measurement over the entire range of visual function, it is possible to establish the reliability of the tests and to discover relations with various measures of job success.

Also, it is possible to apply the principle of the selection ratio for the selection of varying proportions of employees at either end of the scale.

Figure 52 shows a cumulative distribution of scores made by 1,895 women employees in an optical factory on the Ortho-Rater depth test. The selectivity of this test at various levels is indicated by the percentage of workers failing at successive levels, over a range of from 22 per cent to 93

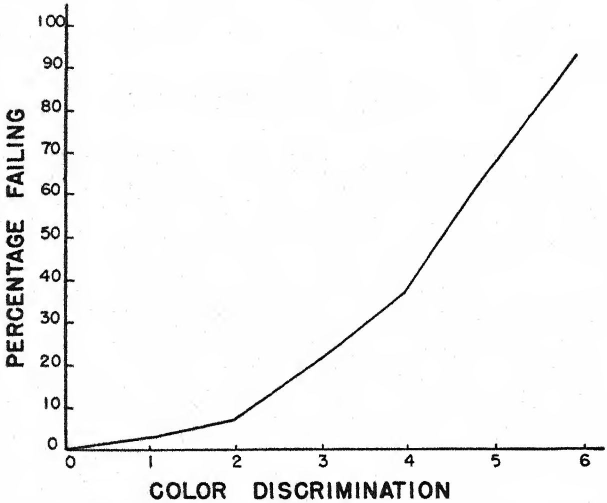


Fig. 53—Percentages of 1,125 employees failing the Ortho-Rater color discrimination test at different levels of difficulty.

per cent. Figure 53 shows, likewise, the scores made by 1,125 employees in a paper mill on the Ortho-Rater color test, with selectivity ranging from 3 per cent to 90 per cent. Each of these functions varies in degree among different employees and can be measured on a continuum with the proper steps or intervals.

All of the visual functions proposed above for use in industry can be measured quantitatively, and should be so measured to give the most satisfactory results in employee placement. Visual "screening tests," using arbitrary stand-

ards or limited in range, are not adequate to show relations between vision and job performance, nor are they satisfactory for placement of employees according to the *amount* of any visual skill they possess.

Measurement for placement

For purposes of allocating employees to the specific jobs in which their chances of success are greatest, and of selecting employees to whom costly training will be offered, it is important to evaluate the different visual requirements of the jobs and to appraise the visual skills of the applicants and employees. The Ortho-Rater tests comprise the most suitable battery of tests, since they were designed and validated for just such purposes. The visual traits as measured in this test battery have been found to be related to safety, successful job performance (production, quality, earnings, scrap savings, and so forth), training time, job tenure, and other criteria of employee desirability. These vision tests make it possible to predict success on the job in terms of their results, for they meet the criteria of good test construction—reliability, standardization of testing conditions and testing method, adequate classification of individual differences throughout the range of human performance, and established relations with various requirements of different jobs. It is possible, with reasonable accuracy, to identify any desired proportion of employees at either end of an acuity scale or of any other of the measures of vision in this battery.

The visual requirements for different jobs cannot be established adequately by observation of the jobs. They can, however, be established by the methods of test validation described in Chapter 3—by either the "new employee" or the "present employee" method. Visual requirements for a given job should be specific not only for the job but also for the workers on the job, or available for it, in a given plant.[6]

[6] Joseph Tiffin and S. E. Wirt, "Determining Visual Standards for Industrial Jobs by Statistical Methods," *Transactions of the American Academy of Ophthalmology and Otolaryngology*, (1945), pp. 72–93.

In other words, vision tests must be validated repeatedly and independently in each situation.

For the purpose of statistical tabulation of Ortho-Rater test results and the establishment of visual requirements for different jobs, Purdue University operates the Occupational Research Center,[7] equipped with modern electric punched-card tabulating equipment and staffed by experienced technicians and psychologists. Results of the studies conducted in this laboratory are reported to the companies participating in the program, together with recommendations for visual requirements and the names of employees who need eye care to meet these standards. Validity studies reported in this chapter are drawn from the files of the Occupational Research Center and from the research that led to the development of the Ortho-Rater.

Measurement for employment

Arbitrary minimum visual requirements for employment in an industrial plant have been increasing in prevalence and comprehensiveness, along with more stringent government standards for license to operate public or private automobiles and airplanes. In the commercial aviation industry, vision standards are comprehensive, severe, and selective for pilots and employees on certain other jobs. The standards are enforced rigorously not only at the time of initial employment but also throughout the entire period of service of these employees. Rigorous enforcement of visual standards has long been a practice among railroads for their trainmen and is rapidly being adopted for drivers of commercial transport trucks and busses. In these industries, the "common carriers," minimum visual and other physical standards are sometimes set and enforced by government regulation in the interests of public safety.

Measurement of vision for this purpose is usually satis-

---

[7] S. E. Wirt, "A Statistical Laboratory for Vision Tests at Purdue University," *Journal of Applied Psychology*, XXX (1946), pp. 354–358.

factory when achieved with simple "screening tests," which measure performance at the lower levels of ability and "screen out" those who score below some arbitrary level on the test. The Snellen letter test is used widely in this way. The Keystone Industrial Visual Service[8] and the Site-Screener[9] are likewise batteries of screening tests useful for this purpose, though they have not been validated in terms of the more critical requirements for placement tests.

A reaction has been voiced against visual screening tests and enforcement of arbitrary minimum visual standards for employment in productive industry.[10] Some workers with substandard vision, especially experienced workers, may be valuable employees on certain jobs if proper precautions are taken for their safety and transportation. Successful attempts to fit seriously handicapped workers into some types of productive jobs in industry have been reported.[11] During periods of emergency production it is necessary to find some job for every possible worker, and even during periods of normal production it is not advisable to segregate a large group to whom employment is denied because of their uncorrectable visual deficiencies, especially if the standards have been set arbitrarily without adequate evidence of the importance of such visual characteristics on the job.

Instances have been found in which arbitrary minimum standards of vision have not only failed to select the better potential employees but have actually selected the poorer ones (see page 206). Such standards also deprive an industry of some of the skillful and experienced workmen who apply for jobs. Even when a correlation exists between visual

---

[8] The Keystone View Company, Meadville, Pennsylvania.

[9] American Optical Company, Southbridge, Massachusetts.

[10] A. C. Snell, "Subnormal Vision and Occupational Aptitude," *New York State Journal of Medicine*, XLI (1941), pp. 1165–1171.

[11] E. B. Merrill, *Occupational Adjustment of the Visually Handicapped.* Publication No. 212 (National Society for the Prevention of Blindness, 1936).

J. Minton, "The One-Eyed Worker," *British Journal of Ophthalmology*, XXIX (1945), pp. 472–476.

performance and efficiency on a certain job, it should not be assumed that every applicant who exceeds a certain standard of performance will be more successful on the job than every applicant who is below the standard. The principle of employee placement in such circumstances is that, other factors being equal or unknown, employees with relatively greater amounts of a certain characteristic are more likely to succeed than others. However, since an employment office often has a limited number of applicants, the standards for employment on any job must be flexible. Moreover, those who are not hired for some specific job may be entirely satisfactory if hired for some other job that has different visual requirements.

Measurement for referral

Visual skills and visual deficiencies can often be improved by professional eye care and treatment. Many employees in industry have a pattern of visual skills that is not adapted to their jobs. Often such employees are unaware of the visual handicap under which they are working. The ophthalmic professions—ophthalmology and optometry—are able to rehabilitate, by means of corrective lenses and/or orthoptic training, a large majority of employees whose vision is not adapted to their jobs.

A visual fact-finding program, such as has already been described, establishes the visual requirements for each job or group of jobs that are similar in visual requirements. Employees who do not have the pattern of visual skills that is necessary or desirable for adequate job performance may be readily identified and referred for professional eye care.[12]

In measuring for referral, employees usually wear their habitual glasses in taking the test, since it would be unwise to refer an employee for eye care if his present glasses solve his

---

[12] S. E. Wirt, "Putting Optometric Service to Work in Industry," *Optometric Weekly*, XXXVII (1946), pp. 1411–1412.

problem. Such measurement should be repeated periodically, on each occasion identifying new cases for professional attention.

Another aspect of measurement for referral to be considered is in connection with the introduction of safety glasses. Employees who do not habitually wear glasses on the job may need something more than plain safety glass before their eyes. Some who do wear glasses may need a change of prescription before safety glasses are made up specially for them. Safety glasses, like any other glasses, should provide lenses that give to the employee good vision for his job. Data summarized on page 209 show that employees who do not have such good vision experience more than their proportionate share of lost time and home case accidents. It is economical to know in advance, if possible, what the lens needs of such employees are so that adequate corrective safety lenses can be supplied. Otherwise, changes in safety glasses might have to be made at the expense either of the company or the employee.

Measurement for compensation purposes

Visual tests given at the time of employment are sometimes useful in compensation cases when it is important to know whether a condition of visual deficiency was caused by injury on the job during an employee's period of service, or whether that condition was present when he was first employed. Several different aspects of visual function have been proposed[13] for appraisal in determining relative loss of competence for earning a living due to injury to the eyes; but in common practice, such consideration is frequently limited to a single visual factor, the ability, measured separately for each eye, to discriminate detail at a standard distance. These tests are always given without glasses.

---

[13] "Report of the Committee on Compensation for Eye Injuries: Appraisal of Loss of Visual Efficiency" (American Medical Association, Atlantic City, N. J., May 26, 1925).

Such tests are usually given in the medical department of a plant as part of the pre-employment physical examination. A widely used test for these purposes is the Snellen letter chart, which consists of several rows of block letters of decreasing size, usually placed at a distance of twenty feet from the subject. A typical Snellen chart is shown in

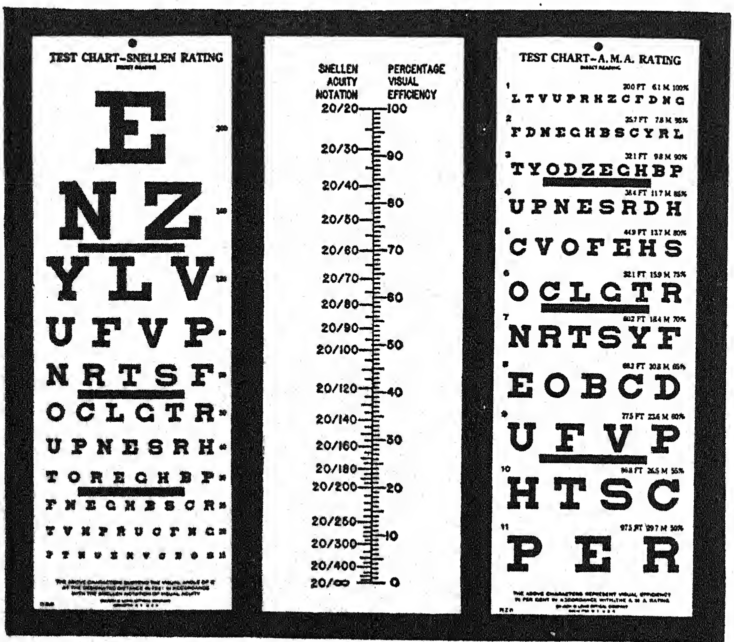


Fig. 54—Snellen and "A.M.A." test charts with scale for converting Snellen scores to per cent of visual efficiency.

Figure 54. The test is administered by determining, separately for each eye, the smallest letters that the subject can read.

The Snellen test is a widely used and effective clinical means of determining visual acuity, particularly in connection with professional eye refraction. Block letters of a certain size and design are generally accepted as a standard of measurement of visual acuity at a distance of 20 feet. The larger the letters at 20 feet that are the smallest readable, the poorer

the visual acuity. The Snellen notation of acuity scores is in the form of a fraction—the smaller the fraction, the poorer the vision. In this fraction the numerator is constant and represents the distance of the test. Thus, visual acuity scored 20/20 is standard. A score of 20/40 means that the subject can read at 20 feet only a letter twice as large as standard, a letter which the "standard eye" can read at 40 feet.

The Snellen acuity designations are not intended to represent fractions of useful vision. The employee who scores 20/40 at a distance of 20 feet is not necessarily handicapped 50 per cent in opportunity to earn a living. In order to simplify interpretation of visual acuity scores and to set up an equitable scale for awarding compensation in proportion to actual incapacity due to eye injury, the American Medical Association has adopted[14] and recommended for use in industry a percentage system of acuity notation, with the distinguishing title of "Visual Efficiency."[15] Where the Snellen test measures acuity in terms of an actual minimum visual angle, the American Medical Association notation interprets this angle in terms of percentage of visual efficiency. The difference between this percentage and 100 per cent is the percentage "loss of vision." A conversion scale for translating acuity scores into the percentage notation is shown in Figure 54. Also shown is the first half of a letter chart that measures visual angles in steps directly equivalent to intervals of 5 per cent on the American Medical Association scale.

This American Medical Association percentage notation is ordinarily not carried above 100 per cent or below 20 per cent, and for compensation purposes in industry no extension of this range is necessary. Acuity above standard is usually considered as only 100 per cent in computing compensation

---

[14] *Ibid.*

[15] Albert C. Snell and Scott Sterling, "The Percentage Evaluation of Macular Vision," *Archives of Ophthalmology*, LIV (1925), pp. 443–461.

awards; and in the absence of records, 100 per cent acuity is assumed prior to the time of an eye injury. Twenty per cent acuity or less is frequently considered as complete lack of acuity or industrial blindness, and is therefore not measured or classified more precisely. These practices vary in different states, and the industrial relations officer should be familiar with the legal practices in his state regarding compensation for eye injury in industry.[16]

In the upper range of acuity scores (standard or above) adequate segregation of differences in acuity is not possible with these letter charts, since the common Snellen charts provide only two or three test levels above standard and the percentage chart provides none. Other limitations of the Snellen letter test have been pointed out,[17] and nonliterate test charts have been developed. These charts eliminate letters and require the subject to identify a spatial pattern, such as a broken ring that may have its open area at the top, bottom, right, or left. In spite of the limitations of both the Snellen and American Medical Association charts, these tests have furnished industry with a convenient and acceptable basis for segregating and classifying cases of substandard acuity in questions of compensation for eye injury. Equivalent information may be obtained, of course, from more elaborate vision tests used in the plant for other purposes.

Vision and Job Proficiency

The importance of vision in industry can be satisfactorily demonstrated only in relation to acceptable industrial criteria, such as hourly production, proportion of work rejected for defective workmanship, supervisors' ratings of employees, employee absences, rate of labor turnover, or some other measurable aspect of employee value to the company.

---

[16] See A. C. Snell, *Medicolegal Ophthalmology* (C. V. Mosby Company, St. Louis, 1940).

[17] C. E. Ferree, and G. Rand, "A New Method of Rating Visual Acuity," *Journal of General Psychology*, XXV (1941), pp. 143–176.

The visual characteristics that correlate with these industrial criteria differ from job to job and cannot consistently be predicted without correlational evidence.

Types of relation

It has been demonstrated on many different jobs that employees who have been classified according to their job

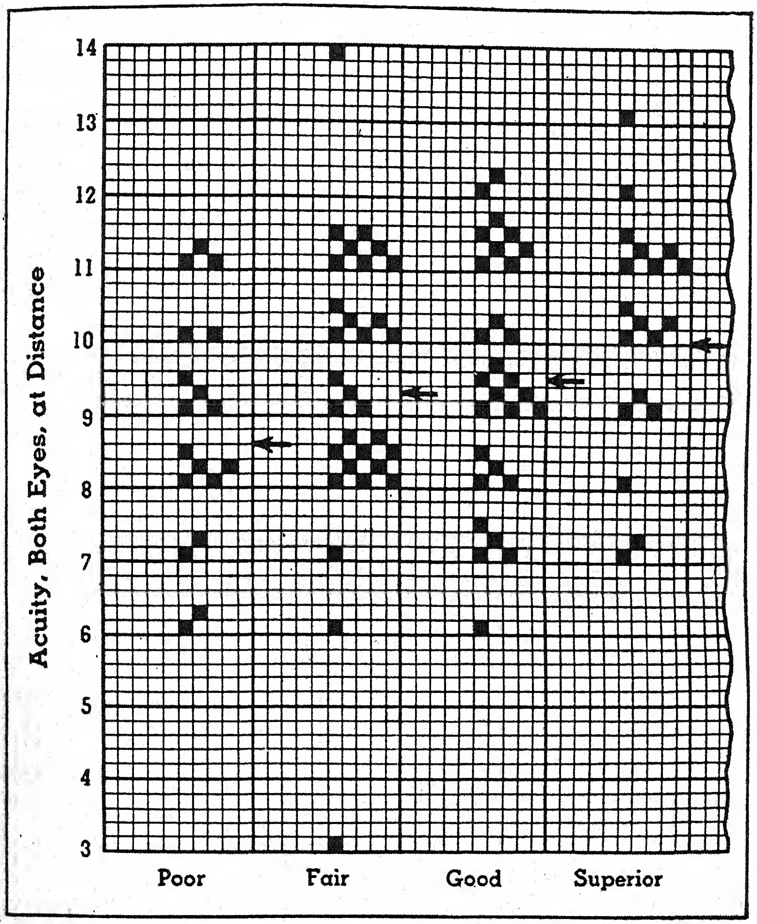


FIG. 55—Distribution of far acuity scores in relation to job performance for 97 milling machine operators.

proficiency do differ with respect to certain visual skills. Those who do the job best usually have certain visual characteristics that are not found as frequently, or to the same degree, among those who do not do it so well. Such a situation was illustrated by Coleman[18] in the case of milling machine operators, as shown in Figure 55. This group of 97

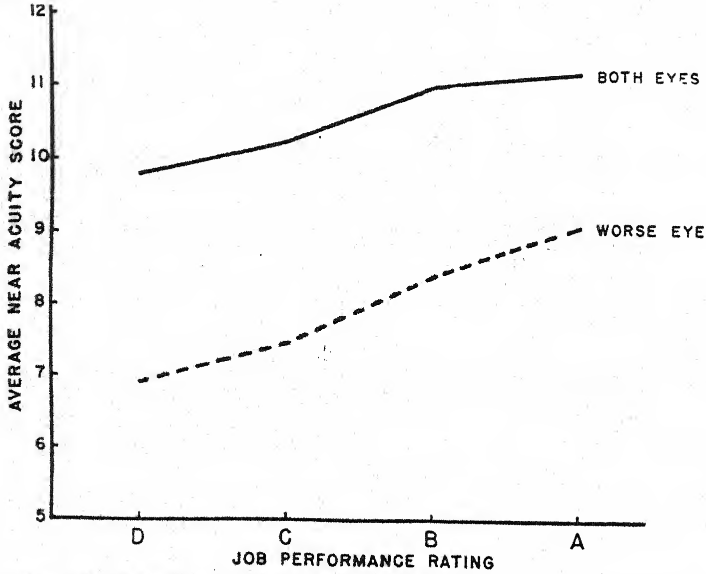


FIG. 56—Differences in near acuity of experienced inspectors with different job performance ratings. $N = 170$.

employees had been classified by their supervisors as poor, fair, good, and superior in job performance. Their scores on distance acuity, for both eyes, on the Ortho-Rater test ranged from 3 to 14. There is considerable overlap in range of acuity of these four classes of employees, but the group averages (shown by arrows) indicate a consistent tendency for the better producers to have higher acuity. Below a score of 9 are 44 per cent of the poor and fair, 31 per cent of the good, and only

[18] J. A. Coleman, "Vision Tests for Better Utilization of Manpower," *Factory Management and Maintenance*, CII (July, 1944), pp. 110–115.

16 per cent of the superior workers. High distance visual acuity is one characteristic more common in the better producers, and is therefore a desirable trait for employees on that job.

Similar trends in near acuity for 170 piston ring inspectors are shown in Figure 56. These inspectors were rated by supervisors as *A* (high), *B*, *C*, and *D* (low) on the basis of production and accuracy. Acuity of both eyes together and of the worse eye separately tend to be consistently higher for

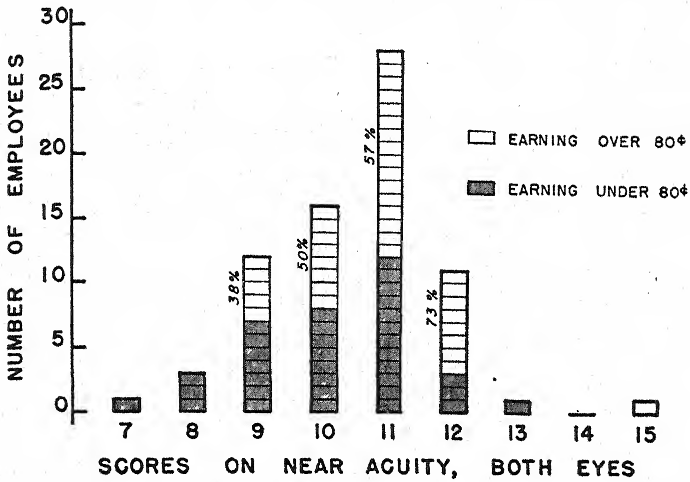


FIG. 57—Average hourly earnings of 73 electric solderers in relation to near visual acuity.

the higher-rated groups. High near acuity, in both eyes and in the worse eye (therefore also in the better eye, or in each eye separately) is a demonstrable requirement for this job. Such a requirement is based on the fact that the better inspectors more frequently score high on acuity, or, to put it another way, the high scores on acuity occur more frequently among the better producers.

The specific relation between any particular vision test and a criterion of success on a given job must be established from a scattergram, like the one on which Figure 55 is based.

The relation may be rectilinear or curvilinear or both in different parts of the test range. A scattergram reveals one or more points in the test range where a minimum requirement can be established that will segregate the largest proportion of superior workers with the smallest proportion of rejections. This point will often be a compromise between the optimum for selectivity of better workers and the maximum of failures that is practical in relation to the current labor market.

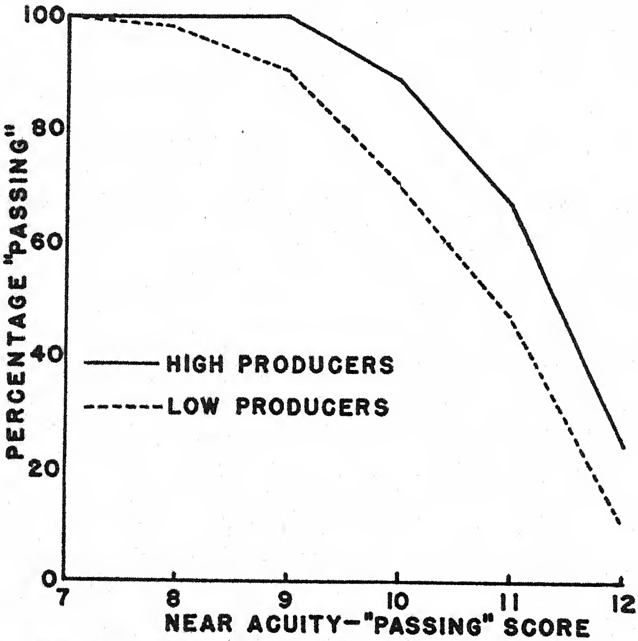


FIG. 58—Selectivity of near acuity test for electric solderers.

Figure 57 shows the relation of near acuity and piece-rate earnings of 73 electric solderers of small metal parts. These solderers scored no lower than 7 on this test. Of those who scored 7 or 8, none was earning the average of 80 cents an hour. Of those who scored 9 and better, increasing percentages earned 80 cents an hour or more. The most select

group of workers comprises those scoring 12 or better, and that score could well be taken as an optimum. So few persons, however, can score so high on the test that this optimum may not be a practical requirement. Figure 58 shows, for different possible minimum requirements, the percentage of the present "better" solderers who could pass the test, and also the percentage of the present "poorer solderers" who could pass. As the minimum requirement is increased from 7 to 12, the proportion of low producers passing the test decreases continuously. But, also, the proportion of high producers passing

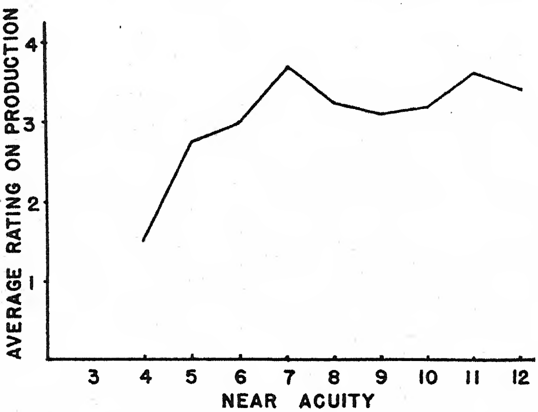


FIG. 59—Rated production of gaugers in relation to near acuity.

the test decreases as the minimum requirement is increased to 10 or more. Since this test is only one of a battery of tests, each of which may likewise be selective for this job, it is expedient to set a minimum requirement on this test that is not of itself too severe and that will eliminate as large a proportion as possible of the poorer workers and as small a proportion as possible of the better workers. In this case, a minimum score of 9 or 10 would be most practical.

It should be remembered, in discussing present employees who "fail" or are "disqualified" on the basis of vision tests, that this result does not imply that they should be taken off

the job. Since visual skills can often be improved through professional eye care, such a classification may mean only that they should be referred for such care.

Figure 59 shows a different type of relation between near acuity and production rating for 177 gaugers. The average rating increases with higher acuity scores, up to 7. Above that point the average rating does not vary systematically. The production averages of all groups scoring 7 or better in

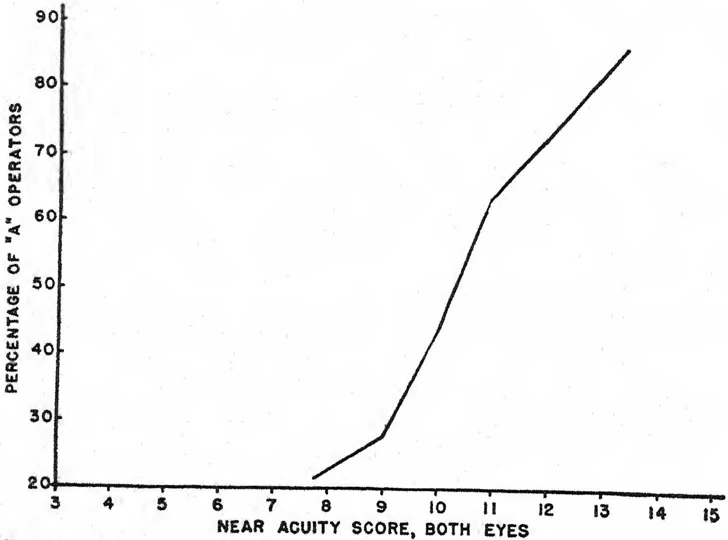


FIG. 60—Rated job performance of radio tube assemblers in relation to near acuity. $N = 225$.

near acuity are higher than are those of any groups scoring below 7. On this test no minimum requirement higher than 7 would improve the discrimination value of the test. It is as though acuity of 7 is adequate to do this job, and better acuity is not needed. This job is therefore in contrast with the previous one mentioned, in which the best acuity for the job seems to be the best that is possible.

Figure 60 shows a positive relation between near acuity and successful work in radio tube assembly. The higher the

score on this test, the higher the proportion of better workers. For this same group, Figure 61 shows a negligible relation between distance acuity and success on the job. Distance acuity is evidently unimportant on this job. In the same manner any one or more of the vision tests comprising a battery may have little or no relation to success on some jobs. It would be uneconomical for purposes of placement to require for this job any minimum of distance acuity,

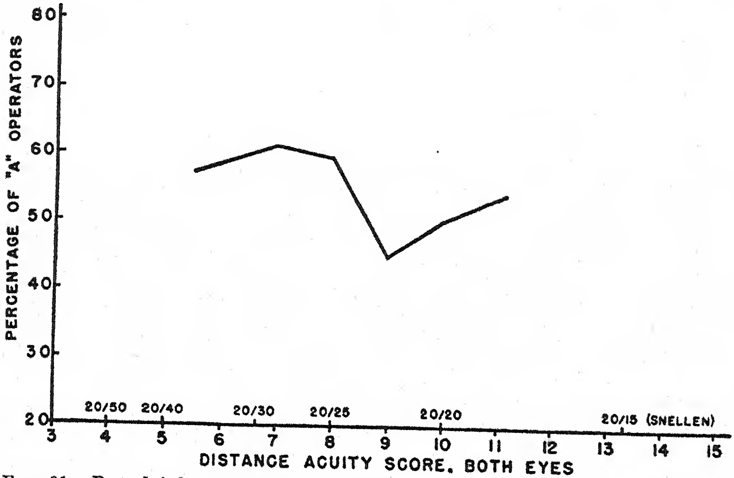


FIG. 61—Rated job performance of radio tube assemblers in relation to far acuity. $N = 225$.

since such a requirement would disqualify some applicants even though it makes no difference on the job.

A negative relation exists between distance acuity and production of hosiery loopers, whose work is at a distance of six to eight inches from their eyes. The better loopers, taken directly from their work for vision tests, do not do as well on distance acuity tests as the poorer workers. Figure 62 shows such a relation for 199 loopers. Those who scored highest on the acuity test, plotted to the right in this graph, were producing on the average 2.5 dozen pairs per hour. Those who scored lowest on distance acuity were producing more than

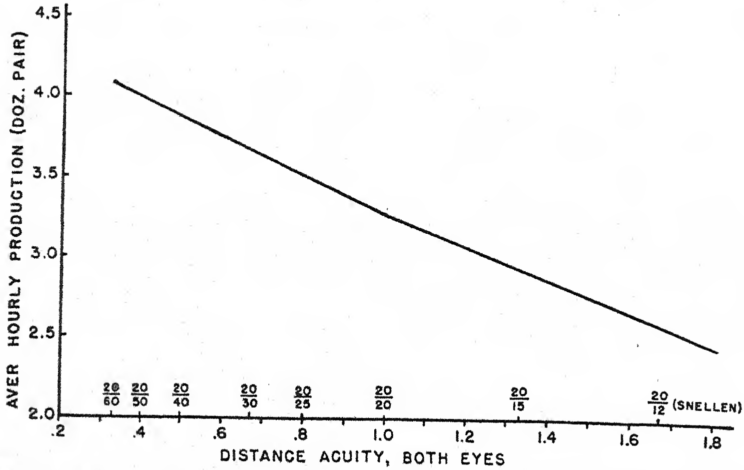


Fig. 62—Hourly production of hosiery loopers in relation to far acuity. $N = 199$.

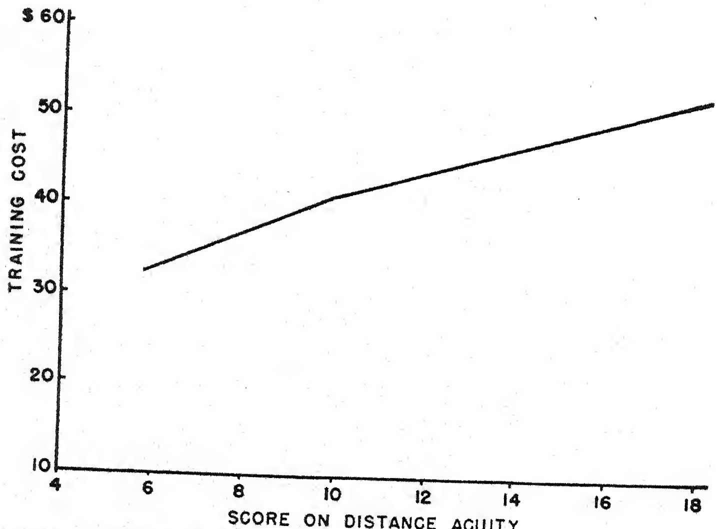


Fig. 63—Training cost in make-up pay for learning loopers in relation to far acuity. $N = 29$.

4.0 dozen pairs per hour. Snellen acuity equivalents are also given in this graph to show that the traditional standard of 20/20 distance acuity or better would, in this case, actually select potentially poorer, rather than better, loopers.

Figure 63 shows a similar relation for 29 looper learners. In this case their far acuity scores at the time of employment are plotted against training cost, which is the total payment supplied by the management to make up minimum guaran-

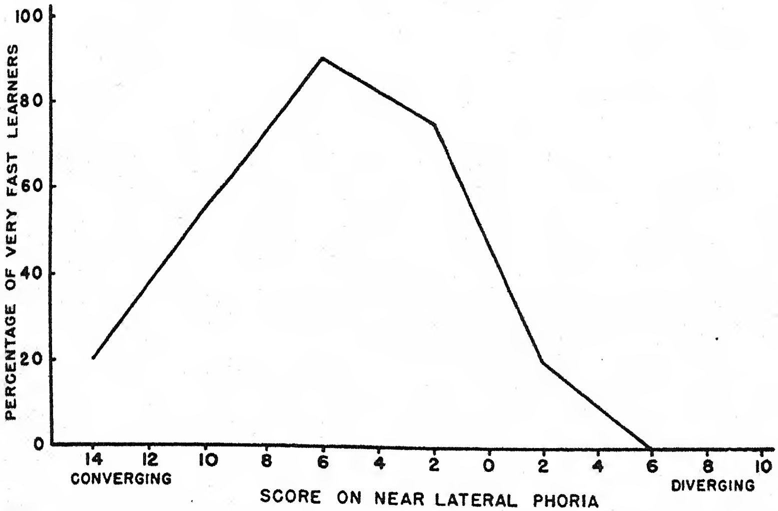


FIG. 64—Speed of learning of loopers in relation to near lateral phoria. $N = 29$.

teed rates, over and above piece-work earnings, until each looper was earning base rate or better when her production was computed at piece-work rates. Those scoring lowest on the acuity test cost on the average about $30.00 each in make-up money, while those scoring highest cost about $50.00.

On acuity, depth, and color tests, the more desirable scores might be expected to be at the high end of the scale. On phoria tests, the more desirable scores might be expected to be in the middle of the scale. Optimum visual scores on phoria tests, however, do vary considerably from job to job. Figure 64 shows, for the same 29 looper learners mentioned above,

test results on near lateral phoria in relation to speed of learning. The phoria test results are here converted to measures of departure from normal convergence for the test distance. The largest proportions of fast learners occur in the range of scores 2 to 8 units to the left of normal, representing greater convergence than normal. Such asymmetry in phoria requirements is not uncommon.

Figure 65 shows a similar asymmetric relation between

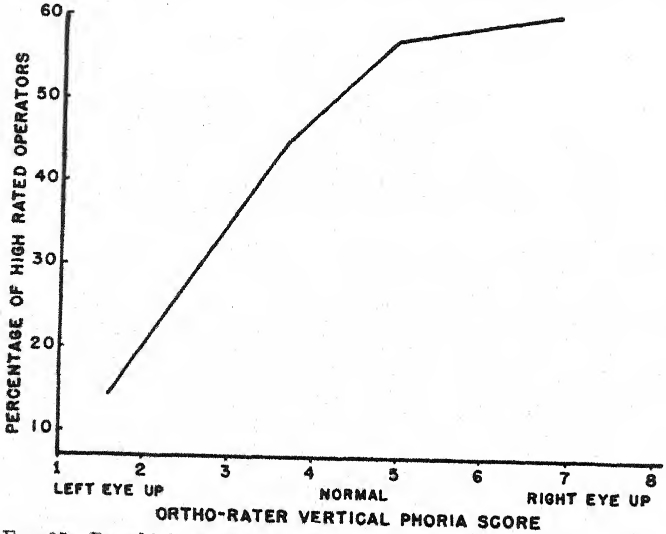


Fig. 65—Rated job performance of milling machine operators in relation to near vertical phoria. $N = 65$.

vertical phoria scores and rated success of 95 milling machine operators. Normal (middle) and slightly abnormal scores at the right end of the scale are acceptable for this job, since they include a large proportion of high-rated operators. On the other hand, abnormal scores in the other direction are undesirable, since they include very few of the high-rated operators. These asymmetric visual requirements are probably associated with unusual postural requirements of the job.

Another asymmetric relation exists between lateral phorias and accident experience in heavy industries. More than

7,000 men employees on all productive and maintenance jobs in a steel mill were tested on a battery of vision tests, and their test scores were considered in relation to whether or not they had had a serious accident in the preceding two years that was unaccountable in terms of some obvious external cause. Figure 66 shows the relation between incidence of such accidents and scores on a far lateral phoria test. The phoria scale is converted here into standard units (prism diopters)

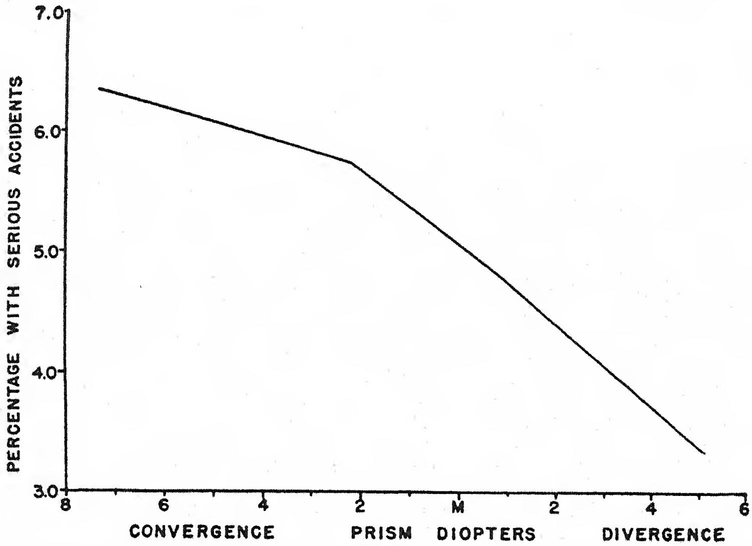


FIG. 66—Relationship between phoria (far) and frequency of serious accident cases.

of deviation from the group mean. Of those cases showing the greatest divergence; about 3 per cent had had serious accidents; of those showing the greatest convergence, over 6 per cent had had accidents.

The opposite relation appeared when the same group was tested for neàr lateral phoria, as is shown in Figure 67. Among those cases showing greatest divergence from the mean, 6 per cent had had accidents; while among those with greatest convergence, 4½ per cent had had accidents. The

accident-free employees score in the range of divergence on the far test and in the range of over-convergence on the near test. This relation was verified by Stump[19] in tests conducted on groups of employees in two plants of a copper and brass company—another heavy metals industry.

This paradox can be resolved by considering the phoria posture as an indirect response to the focus stimuli in the phoria tests. A change in focus demand from far to near

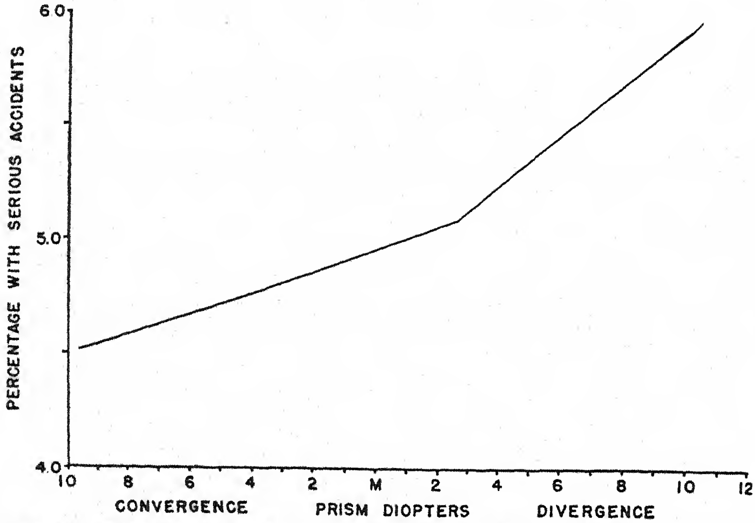


Fig. 67—Relationship between phoria (near) and frequency of serious accident cases.

induces a change of convergence posture in the phoria tests. Subjects with the greatest change in convergence from the far test to the near (divergence at far, convergence at near) are most responsive to a change in focus stimulus and are also more frequently accident free. Figure 68 shows the accident experience of these same workers in a steel mill as it is related

[19] N. F. Stump, "Spotting Accident-Prone Workers by Vision Tests," *Factory Management and Maintenance*, June 1945, pp. 109–112. See also "A Statistical Study of Visual Functions and Safety," *Journal of Applied Psychology*, XXIX (1945), pp. 467–470.

to a composite of both phoria test scores—the actual change in convergence from far test to near test. Of those with the greatest increase in convergence on the near test, 4 per cent had had accidents, while among those with a negative increase in convergence, 9 per cent had had accidents. Stump made similar computations in terms simply of far phoria score minus near phoria score plus a constant.

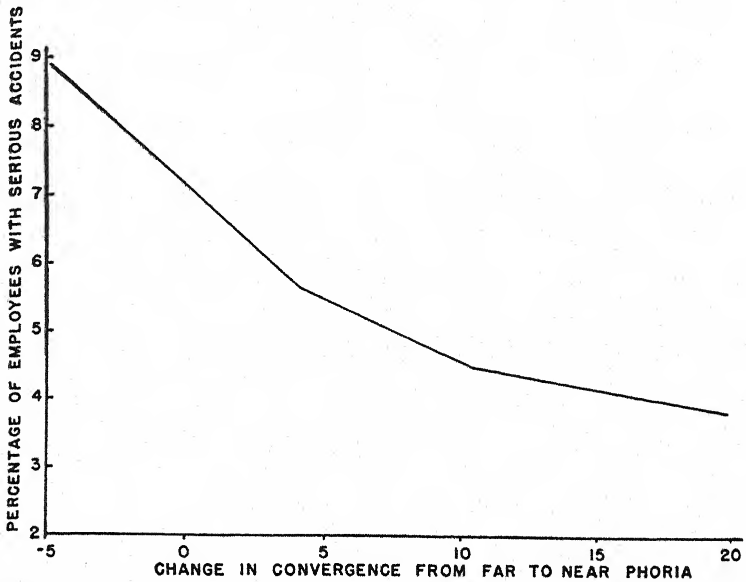


FIG. 68—Net relation between lateral phorias, far and near and incidence of serious accidents in a steel mill. $N = 8503$.

It is evident from the foregoing examples that specific vision tests, properly constructed and possessing adequate reliability, have different relations with successful performance on different jobs. Some of these relations might have been anticipated, but others most likely would not have been inferred from observation of the jobs. It is also evident that such relations between test results and various measures of employee success on the job can be discovered. It is quite practicable, from a study of these relations, to establish

minimum and/or optimum visual requirements on specific tests for a given job. The pattern of visual requirements established by such a factual appraisal of visual test results represents the pattern of visual skills that is desirable for that job. This pattern constitutes "good vision" for the job.

Patterns of visual requirements

Each test in the battery of vision tests is separately analyzed in relation to job success. Some of the tests may

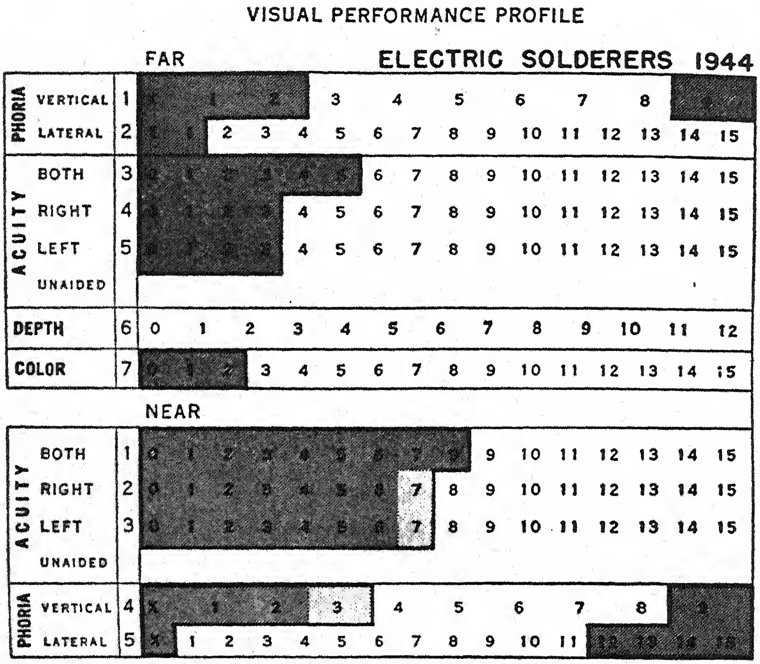


FIG. 69—Pattern of visual requirements for electric solderers.

show a relation to success on a particular job, others may not. On each related test, a range of scores is designated as optimum for the job. The total pattern of these visual requirements represents the pattern of visual skills that is predictive of success on this job. Such a pattern, or profile, of visual

requirements for the job of electric soldering mentioned above is illustrated in Figure 69.

In this profile the blacked-out areas represent undesirable scores in the prediction of success on this job. The cross-hatched areas represent undesirable scores for maximum success.[20]

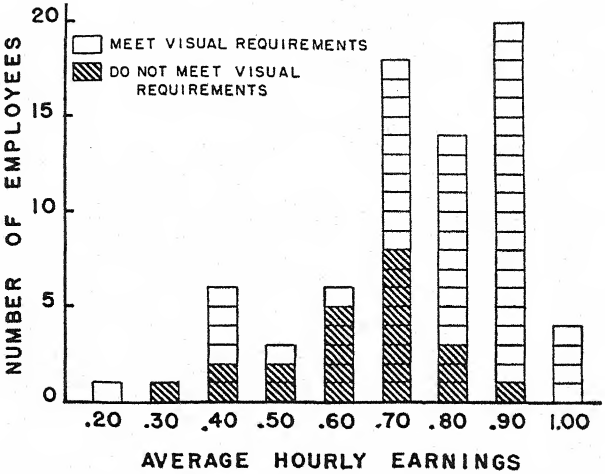


FIG. 70—Hourly earnings of electric solderers who do or do not meet the visual requirements for the job. $N = 70$.

All present employees on this job were classified as to whether they did or did not meet all the visual requirements of the battery of tests. In terms of average hourly earnings, those who did meet them were producing more than the others. Figure 70 shows this relation between earnings and total visual requirements. Of those solderers earning below 70 cents an hour, 41 per cent met all the visual requirements

[20] This and all other visual profiles discussed in this chapter were established in the Occupational Research Center at Purdue University and are based on data sent in for analysis by various industrial plants using the Ortho-Rater. The Purdue University Occupational Research Center, which is sponsored and subsidized by the Bausch and Lomb Optical Company, is a statistical research unit concerned primarily with the analysis of visual and other personnel test data in relation to the requirements of industrial jobs.

(shown by clear spaces). Of those earning 70 to 89 cents an hour, 66 per cent met all the requirements; and of those earning 90 cents or more, 96 per cent met all the requirements. Of those meeting the visual requirements, 67 per cent earned over 80 cents an hour; but of those who did not meet the requirements, only 18 per cent earned over 80 cents an hour. Those who met the requirements produced 12 per cent more than the visually unqualified group.

This pattern of visual requirements also differentiated employees in terms of quality of work, as shown in Figure 71. Twenty per cent of the work of those employees who did not

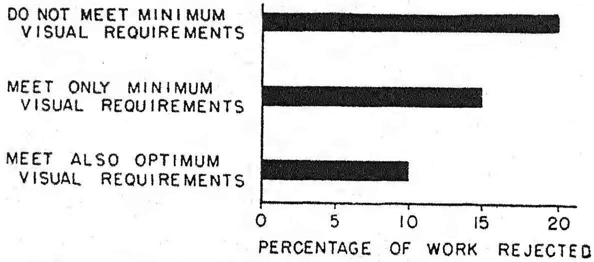


FIG. 71—Quality of work in relation to vision of experienced electric solderers. $N = 70$.

meet minimum visual requirements was rejected by inspectors and returned to the solderers for repair or scrap. Of the work of those who met the minimum visual requirements, 15 per cent was rejected. For those who could also meet a still more severe pattern of visual requirements (established on a review of this job a year later), the rejection rate was only 10 per cent —reinspection and repair costs were only half as great as for the visually unqualified group.

Also this visual pattern differentiated beginners in terms of learning time. Forty per cent of the new employees put on this job who did not meet all the visual requirements were earning their base rate in ten weeks time; but 75 per cent of those who did meet all the visual requirements were earning base rate or better in ten weeks time. The company paid on

an average $30.00 less in minimum wage make-up money to each of those who were visually qualified than to each of those not qualified. Of those who were visually qualified, more were still on the job a year later; and they had had somewhat fewer absences than those who were not visually qualified.

In this pattern of visual requirements, near vision is more important than far vision. In the re-study of this same job

VISUAL PERFORMANCE PROFILE

LOADING & PUSH TRUCK

FAR

PHORIA	VERTICAL	1	X	1	2	3	4	5	6	7	8	9						
	LATERAL	2	X	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
ACUITY	BOTH	3	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
	RIGHT	4	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
	LEFT	5	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
	UNAIDED																	
DEPTH		6	0	1	2	3	4	5	6	7	8	9	10	11	12			
COLOR		7	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15

NEAR

ACUITY	BOTH	1	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
	RIGHT	2	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
	LEFT	3	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
	UNAIDED																	
PHORIA	VERTICAL	4	X	1	2	3	4	5	6	7	8	9						
	LATERAL	5	X	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15

Fig. 72—Pattern of visual requirements for loading and push truck operators.

a year later,[21] the far acuity requirements were eliminated completely.

A radically different pattern of visual skills, that desirable for push truck operators, is shown in Figure 72. Here only a modest degree of distance vision is important for success on

[21] S. E. Wirt, "Eye Care Increases Earnings of Electric Solderers," *Optometric Weekly*, XXXVII (1946), pp. 1091–1093, 1104.

the job. Another type of visual skills profile, for workers operating milling and profiling machines, is shown in Figure 73. Here both far and near vision are important. On still another job, gray inspection in a hosiery mill, good one-eyed vision is quite acceptable, as shown in Figure 74.

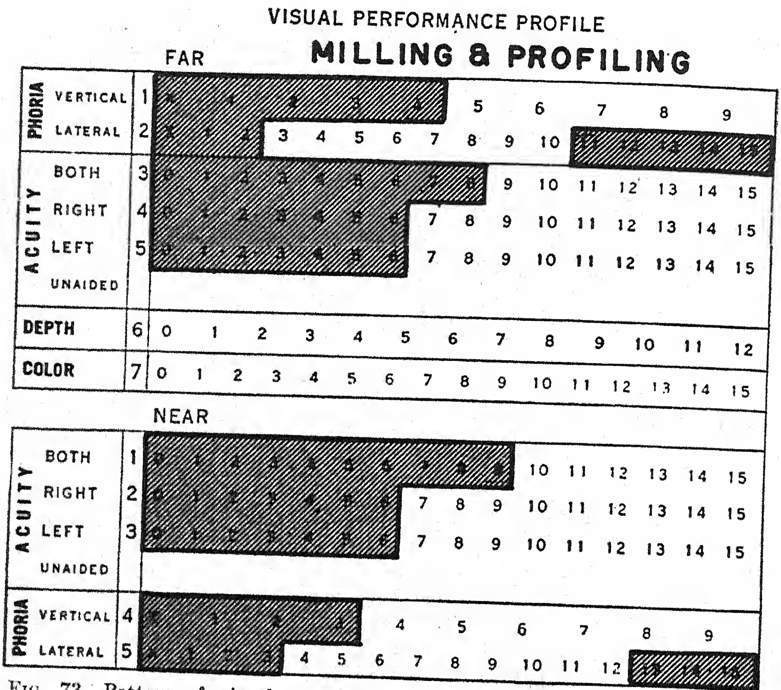


FIG. 73—Pattern of visual requirements for milling and profiling machine operators.

Such different patterns of "good vision" have been established for many jobs in various plants. Similar operations have similar visual requirements. Similarity of job title in different plants, however, does not guarantee similar operations or similar visual requirements. And even on similar operations, the severity of visual requirements may vary from one plant to another, depending on the general level of visual skills in the industrial population. But

for a given job in a given industrial population, practical visual requirements can be established by the method here illustrated.

Not often is there available for test validation such a wide array of criteria as in the study on electric solderers. The visual requirements established for that job discriminated

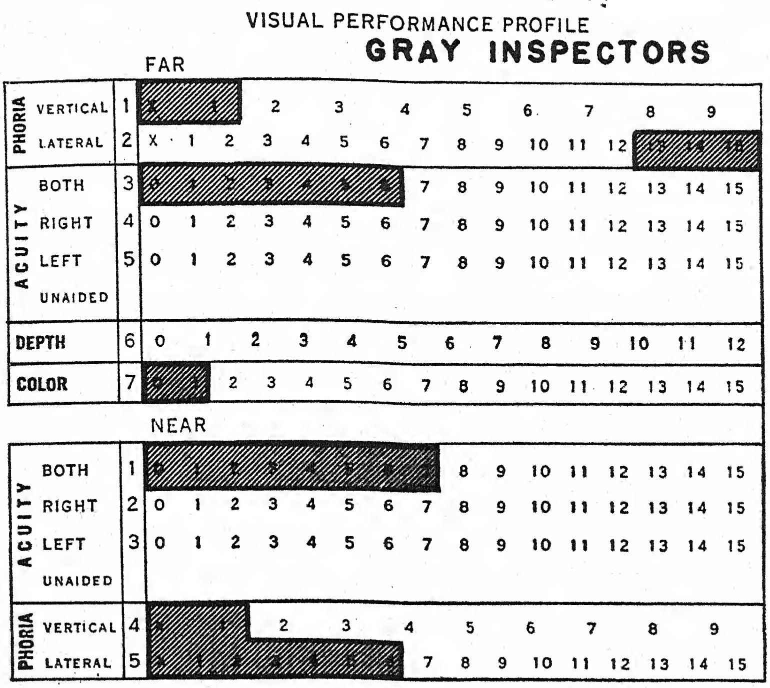


FIG. 74—Pattern of visual requirements for hosiery inspectors.

among workers in terms of production, earnings, quality, ratings, accidents, absences, tenure, learning time, and training cost. Occasionally a pattern of visual requirements will correlate positively with one measure of job success and negatively with another. In one such case, for a job of assembling small electrical parts, minimum requirements on several vision tests correlated positively with quality of work and negatively with quantity. Figure 75 shows these results.

In the upper part of the figure, the greater length of the black bars shows a higher average hourly earning (quantity of work) for those failing on each of four vision tests than for those passing. In the lower part of this figure, the greater length of the first black bar shows that a higher rating for quality of work, also, was given to those failing the far acuity test. But the three longer white bars show that on the other three tests work of higher quality was done by the passing

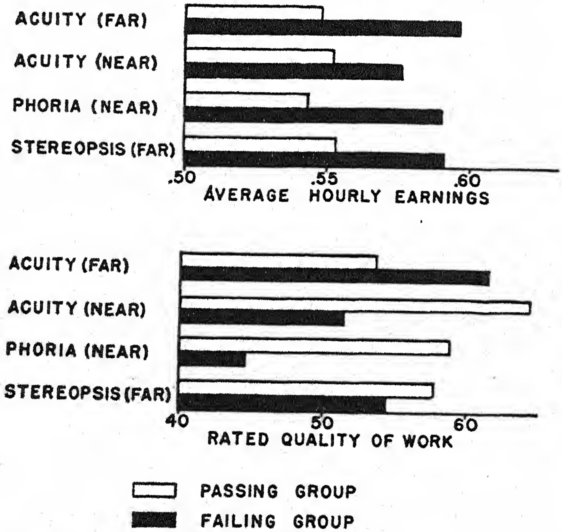


FIG. 75—Differences in quantity and quality of work among 33 electrical assemblers who passed or failed 4 vision tests.

group. Except for far acuity, ability to pass the vision tests correlates with lower production and better quality of work. Ability to pass the far acuity test correlates with lower production and poorer quality. Ayers[22] reported similar results from an investigation conducted among inspectors of cones of rayon thread. In such situations the different criteria are

[22] A. W. Ayers, "A Comparison of Certain Visual Factors with the Efficiency of Textile Inspectors," *Journal of Applied Psychology*, XXVI (1942), pp. 812–827.

inversely related to each other. Management must then decide what measure of job success best represents the kind of workers that is wanted, and the pattern of desirable visual skills can then be established on the basis of this criterion. More frequently the criteria of job success are interrelated, and for most measures the same pattern of visual skills discriminates between better and poorer workers. For example, a pattern of visual requirements for general job efficiency, based on supervisors ratings, was established for paper machine operators. This same pattern correlated with accident experience, as shown in Figure 76.[23] Of those

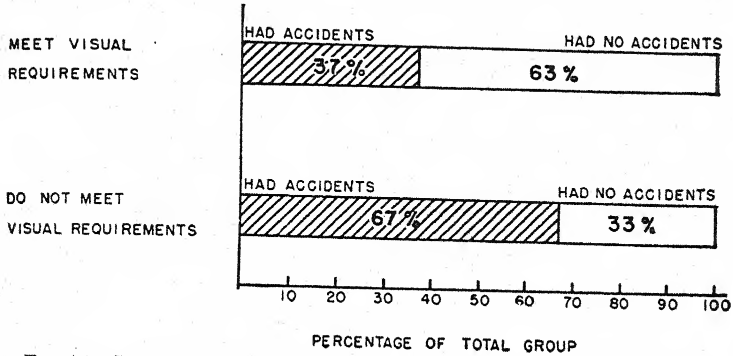


FIG. 76—Incidence of accidents serious enough to require medical attention among paper machine operators who do or do not meet the visual requirements for the job.

meeting all visual requirements for efficiency, 37 per cent had had accidents serious enough to require medical attention in the previous year, while among those who did not meet all the visual requirements, 67 per cent had had accidents.

Changes in Vision with Age and Job Experience

Visual skills change with age, as well as with continued experience and visual application in certain work. Entirely apart from the influence of occupational trends and loss of

---

[23] S. E. Wirt and H. H. Leedke, "Skillful Eyes Prevent Accidents," *Annual News Letter*, National Safety Council, Industrial Nursing Section, Nov. 1945, pp. 10–12.

vision from injury, visual skills show almost universal deterioration with advancing age. Consequently, employees who have at any time all the desirable visual characteristics for their jobs will not likely maintain those skills indefinitely without professional eye care. Fortunately, most of the natural losses of visual efficiency with increasing age can be adequately compensated for by means of professional eye care and optical aids.

Recent research has increased our knowledge of the scope of these changes in maturity with respect to differences between men and women, differences among different industrial populations, and the effect of optical aids. Maintenance of visual skills involves another aspect of vision testing—the periodic re-test to disclose losses of visual skills that might be repaired. The facts concerning visual changes with age indicate the ages at which such changes are most pronounced, during which years periodic re-tests are most important. Since the proportion of older employees in industry has been increasing for at least a half century, these changes in vision must be of greater concern in industry than ever before.

Visual acuity and age

One change in vision with age is common knowledge—the change that makes it necessary for older persons to use extra lens-power to see clearly at close distances. Figure 77 shows composite figures, representing over 11,000 cases from a number of industrial plants, on near acuity, both eyes, together for men and women who were using glasses for near seeing and separately for men and women who did not use glasses. All three classes are sub-divided into five-year age groups. Among those without glasses, there is loss of near acuity for some before age 35, and serious loss by age 45. Loss in this particular visual skill comes in middle age—the age range that embraces most of the key men in any industry. By age 55 this transition is pretty well completed; further losses are slight.

Up to age 45 women without glasses have on the average less keen near vision than men, and after age 45 have retained somewhat keener near vision. This difference between men and women is substantiated in a study by Collins,[24] summarized by Wirt,[25] showing that women enter upon and complete this transition several years earlier than men.

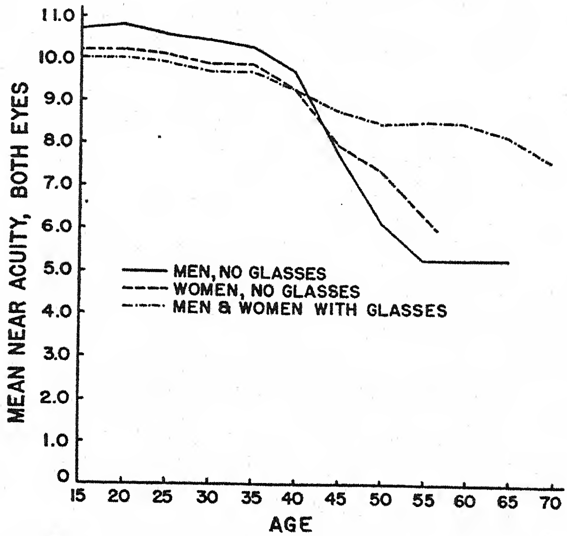


FIG. 77—Changes in near visual acuity with age for men and women who do and do not wear glasses.

Men and women who wear glasses for near seeing, even though the glasses may no longer be adequate for best vision or may have been obtained without benefit of professional eye care, tend to maintain a more constant level of near acuity (and a much higher level) after age 45 than those without glasses. This evidence indicates that with adequate professional eye care and optical aids, losses of near vision can largely be eliminated.

---

[24] S. D. Collins and E. H. Pennell, "The Use of the Logistic Curve to Represent the Prevalence of Defective Vision of Persons of Specific Ages Above 30 Years," *Human Biology*, VII (1935), pp. 257–266.

[25] S. E. Wirt, "Industry Faces Presbyopia," *Optometric Weekly*, XXXV (1944), pp. 653–654.

Figure 78 shows differences in far acuity between employees who do and do not use spectacles for distance or general vision. Below the age of fifty those wearing spectacles have lower average acuity than those who do not use spectacles. Probably those in this age range who are using glasses are the ones who most need them to bring acuity up to average, and

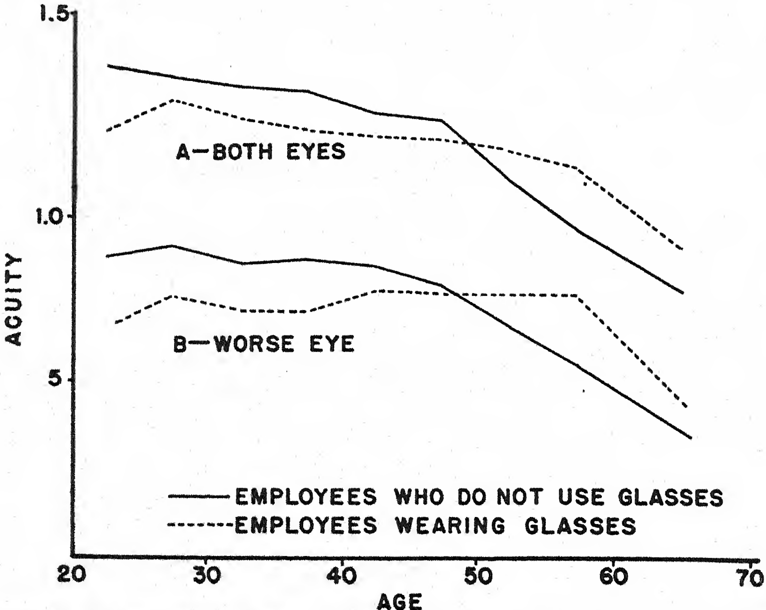


FIG. 78—Differences in far acuity between employees who do and do not use spectacles.

those who do without glasses most frequently have average or superior acuity. After the age of fifty, however, those who are already accustomed to spectacles maintain their vision by occasional change of glasses, but those who have done without spectacles suffer considerable loss before they obtain help. Figure 79[26] shows much less rapid loss of distance

[26] Jesse C. Rupe, "Variations of Visual Skills with Sex and Age," A Thesis submitted to the faculty of Purdue University in partial fulfillment of the requirements for the degree of Master of Science, 1946, p. 19.

acuity for men and women in another plant where for a number of years professional eye care had been provided for employees. In this graph women show lower average acuity than men, in spite of the fact that a larger proportion of them use eyeglasses. A reasonable speculation in explaining this difference is that women in a plant are more frequently employed on jobs requiring visual adaptation to prolonged near vision, and that, in addition, men more frequently have

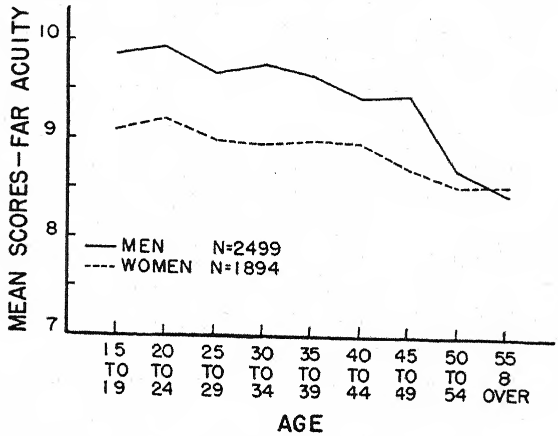


FIG. 79—Changes in far visual acuity with age for men and women employees in a precision industry.

outside interests that exercise distant vision—such as driving and sports.

Depth perception and age

Figure 80 shows differences in average depth perception scores with increasing age for three industrial populations—2,200 men in heavy industry (paper, steel, shipbuilding), and 1,800 women and 2,400 men, separately, in light precision work. All those employees who used glasses on the job were tested wearing them.

Men and women in the same light industry show somewhat different patterns of change in depth perception with age. This difference is probably associated with the fact that for

the most part men and women perform different jobs in the plant—most of the heavier and physically more active jobs being done by the men. The men tend to show a higher development of depth perception during middle age—the only example of a visual skill increasing with age. This trend is more pronounced among men in the heavy industries. It would seem that physically active work is associated with a greater development of depth perception, which does not

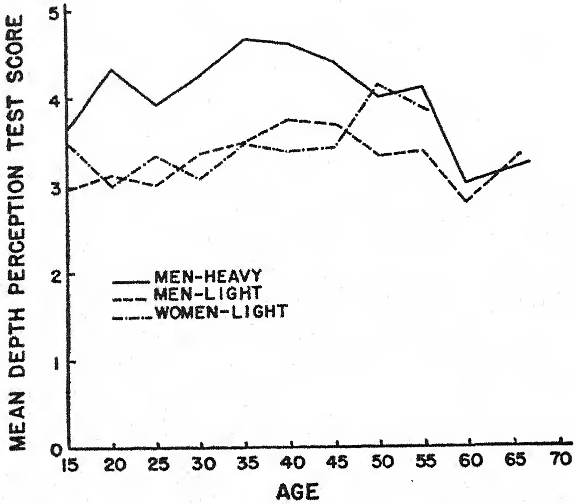


FIG. 80—Changes in depth perception with age for different groups of industrial employees.

reach its peak of development until age 35. After age 50 loss of depth perception accompanies loss of distance acuity. Since this visual skill develops with experience in certain types of work, then presumably it could be developed earlier, by training, for those jobs where it is needed.

Color discrimination and age

Figure 81 shows mean scores on the Ortho-Rater color test for large samples of men and women from a number of industries, classified by age. The test is similar in form to the

Ishihara Color Perception Test, and probably measures the same visual function. Women tend to average very slightly higher than men in color test scores, and after age 45 both sexes show a loss in color vision. In an earlier report[27] based on a study of a pre-war population of men in a steel mill, it was shown that decreases in color vision began by age

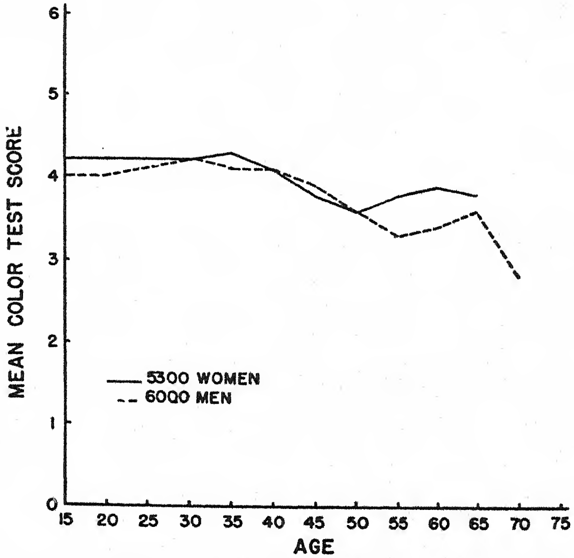


FIG. 81—Changes in color discrimination with age for men and women.

25. Both studies agree that color vision deteriorates with advanced age. On jobs requiring a combination of color discrimination and long training, it therefore becomes very important to select trainees with a great initial amount of color-vision ability.

Some changes in vision with job experience

Experience on a particular job often tends to develop in employees some of the visual characteristics that are desirable

[27] Joseph Tiffin and Hedwig S. Kuhn, "Color Discrimination in Industry," *Archives of Ophthalmology*, XXVIII (1942), pp. 851–859.

on the job and some that may not be desirable. These changes may be in the same direction as the changes with age or in the opposite direction. Loopers in hosiery mills tend to develop exceptional amplitude in focus range for very near objects and to maintain this facility in spite of the general tendency to lose it with increasing age. Exceptional visual performance that characterizes experienced employees on a particular job is in part the result of practice and incidental

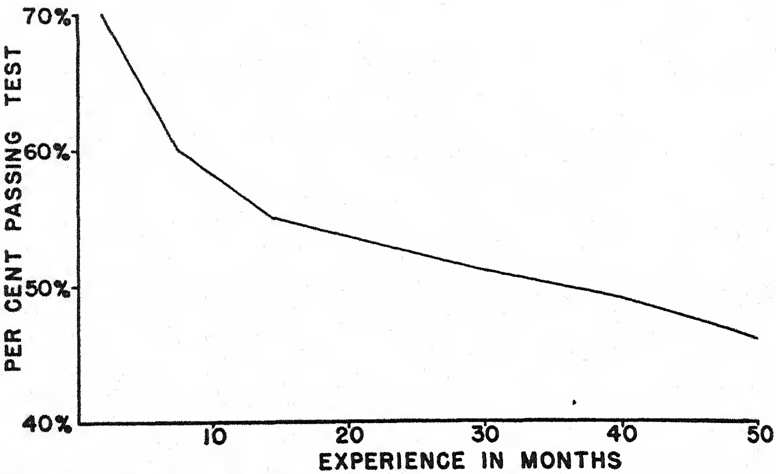


Fig. 82—Changes in distance acuity with increasing experience on the job of looping hosiery. Curve based on 206 hosiery loopers.

training on the job and in part the result of a selection process, entirely apart from any formal selection with respect to vision. Those employees who have or can easily develop the desirable qualifications for the job tend to remain on the job and accumulate experience. Those who do not have or cannot easily develop these qualifications tend to drop out or to be transferred to other jobs before they acquire much experience. Consequently, the relationship between visual characteristics and experience is augmented by this selection process that operates in the same general direction as experience. The

results of such selective and adaptive processes have been investigated for jobs requiring close vision.[28]

The first year of looping hosiery is a period during which marked changes in visual characteristics occur. Figure 82 shows the relation between experience and distance acuity for 206 loopers. Distance acuity tends to decrease with experience, and this relation holds whether employees wear correc-

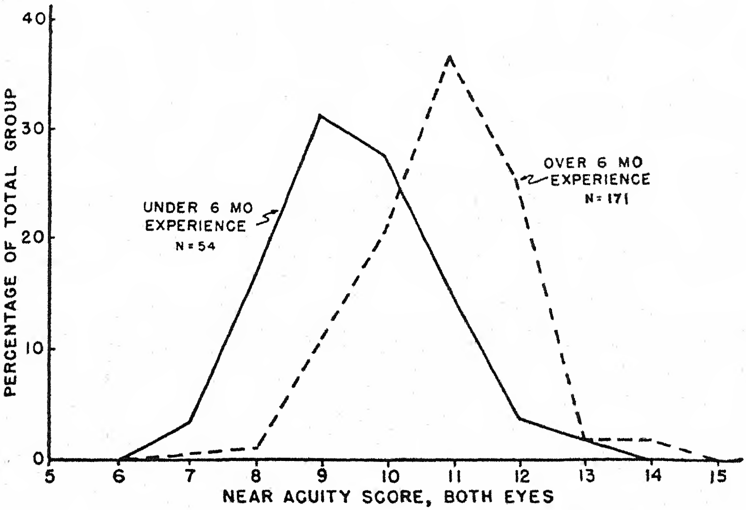


FIG. 83—Differences in near acuity of radio tube assemblers with greater and less job experience.

tive spectacles or not. This change may be interpreted as a shift in visual skills from far to near, for these loopers showed exceptionally high development of near vision.

Figure 83 shows another relation between vision and job experience—in this case between near acuity and experience on the job of radio tube assembly. Fifty-four assemblers with less than six months experience scored from 7 to 13, the

[28] W. S. Duke-Elder, "An Investigation Into the Effect Upon the Eyes of Occupations Involving Close Work," *British Journal of Ophthalmology*, XIV (1930), pp. 610–620. Also see M. Luckiesh and F. K. Moss, "Functional Adaptation to Near-Vision," *Journal of Experimental Psychology*, XXVI (1940), pp. 352–356.

greatest number scoring 9. Among 71 assemblers with more than six months experience, scores ranged from 7 to 14, with the greatest number at 11. A minority of the inexperienced workers and a majority of experienced workers scored higher than 10 on this test. The greater incidence of higher near acuity among experienced assemblers indicates either (1) a natural elimination from this group of many beginners who did not have high near acuity, or (2) a development of greater near acuity with increasing job experience. The fact that about an equal proportion of both groups were wearing eyeglasses indicates that neither group had any marked advantage in eye care.

If the higher near acuity of the experienced group is due to natural selection, it is then important to select new employees with adequate near acuity so that more of them will remain on the job. If the higher acuity is due to learning on the job, then the development of near acuity is part of the learning required on the job. Selection of beginners with adequate near acuity could be expected to reduce the amount of learning, and therefore the training time, required on this job.

Job differences in vision

Visual conditions that characterize employees on several jobs in a sheet and tin mill indicate the effects of experience and informal selection. Formal selection standards for initial employment on these jobs included only acuity and, for some jobs, color vision. But, as shown in discussing age trends, even these differences in vision due to differential selection for initial employment on different jobs are rapidly eliminated by age trends and also by transfer from one job to another. Certain job factors, however, result in training and further selection on the job. Table 19 gives the percentages of employees on different payroll jobs who met certain plant-wide visual standards. This table also shows, for comparison, the percentage of all employees in the plant who met these

standards. Small differences between groups are of no significance, but large differences indicate real variations from the plant average and indicate real job differences. Those figures in Table 19 that differ by a statistically significant amount from the plant average are marked. Every job row and every vision column shows one or more significant differences.

TABLE 19

PERCENTAGE OF EMPLOYEES ON DIFFERENT JOBS WHO PASSED VARIOUS TESTS OF VISION

Employees	Approximate N	Acuity, Far, Both Eyes	Acuity, Far, Worse Eye	Depth Perception	Color Vision	Phoria, Far	Phoria, Near	Acuity Near
1. Foremen...	320	95.9	77.8	85.1*	66.5*	57.0	63.7	81.4*
2. Clerks.....	590	94.1	73.7	79.5	79.2*	63.1	69.4*	88.9*
3. Electricians	430	94.5	68.4†	86.4*	59.6	57.9	59.6	61.7
4. Machinists.	380	93.7	73.3	92.4*	60.0	60.9	66.2	65.6*
5. Cranemen..	250	97.6*	82.1*	81.9	57.0	61.3	66.2	52.2
6. Hookers....	250	94.7	75.0	81.0	59.1	71.1*	67.9	59.8
7. Laborers...	710	92.3	72.2	74.5†	52.7	62.9	66.0	52.4†
Plant Total...	7000	94.0	73.1	79.2	55.3	61.1	63.6	61.0

* Significantly high. † Significantly low.

Note: These figures are corrected for differences in age on the different job groups.

If employees on different jobs are drawn from different populations, differences from job to job may be due to differences in the source of the employees. Table 20 shows separately for white men, Negro men, and white women in a large plant the percentages passing a test of distance acuity and a test of red-green color discrimination. The women were best on color vision, the Negro men best on acuity. Since Negroes and women in this plant were only on certain types of jobs, these differences between races and sexes might be mistaken for characteristic job differences. Also differences in age have not been ruled out in this comparison.

In this same plant, watchmen and police in the plant protection department showed surprisingly poor scores on many of the vision tests, especially distance acuity. One

reason was their average age, which was considerably more than the plant average. Another possible reason was that in the past veteran employees who could no longer be kept on productive jobs were transferred to the plant protection department. These extraneous selective factors will often obscure job differences in vision or create spurious differences unless they are carefully controlled in statistical studies of job differences.

TABLE 20
PERCENTAGE OF NEGRO MEN, WHITE MEN, AND WHITE WOMEN EMPLOYEES PASSING TWO VISION TESTS

EMPLOYEES	PERCENTAGE PASSING TESTS	
	Color Vision	Distance Acuity
White Women	74.5	76.0
White Men	60.4	76.1
Negro Men	40.7	80.7

Visual Improvement

Unlike most other aptitudes and traits that may be considered for employment, visual skills can be improved or adapted to specific demands by means of professional eye examination and eye care. Applicants and employees who do not have the visual skills that are important for their jobs can usually be given the kind of vision they need. This possibility is important when labor is scarce and the available employees cannot all qualify visually for the jobs for which they are needed. It is also important to realize, in the light of changes in vision with age, that the visual efficiency of most experienced workers can be maintained at an adequate level.

Use of glasses

Use of eyeglasses on the job is an imperfect indication of the extent to which employees have taken advantage of

professional eye care. In some cases glasses are bought over a counter without benefit of professional advice. In other cases the glasses are obsolete, or might have been prescribed for other activities. For example, in a certain steel mill about 20 per cent of the men over fifty years of age carried in their pockets glasses that they used only when they needed to see better at close range—in some cases these glasses were used rarely or not at all on the job. Many of them were "store glasses," not fitted to the individual needs

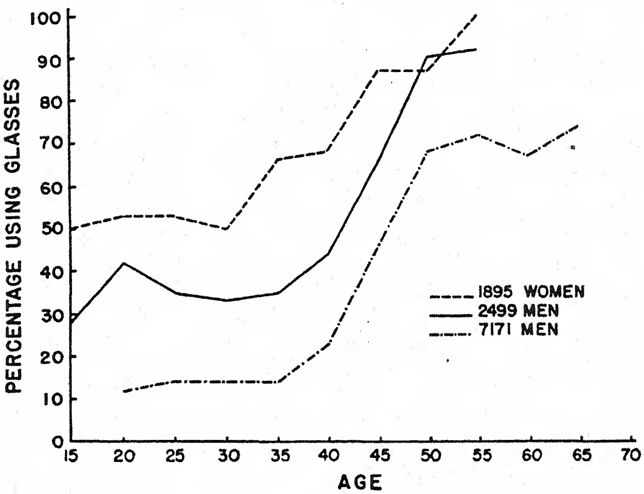


Fig. 84—Use of glasses at different ages for different industrial populations.

of their wearers; 75 per cent of them did not even enable their users to see clearly at close range. However, 50 per cent of the men over age 50 wore glasses regularly on the job, most of these glasses having been fitted by professional eye men. Only 20 per cent of these glasses did not enable their users to see clearly at close range.

Use of glasses is at least an indication of a concern for vision and a search for more efficient or more comfortable seeing. Figure 84 shows, for three industrial groups, the proportion of employees of different ages who used any kind

of glasses at work. The lowest curve represents 7,171 men on productive and maintenance jobs in a sheet steel mill. The upper two curves represent 1,895 women and 2,499 men in a precision instrument plant. At all ages, the men in the precision industry used glasses more frequently than men in the heavy industry, and women used them more frequently than either. Exacting visual demands, and therefore greater need of eye care, are more common in the precision industry, and particularly for women's jobs. Further, adequate eye care was available at the plant to all employees in the precision work, and their social and educational background was such as to lead them to take advantage of all health and professional services. Whether employees take advantage of eye care depends on the availability of eye care, the visual demands of their jobs, their socio-economic status, and their age.

Not half the men in the heavy industry used glasses before the age of fifty—late in life in relation to the losses of vision with age as shown in the previous section. In contrast, over half of the men in the precision work used glasses by age forty-five, and over half the women used them by age thirty-five.

Eye care improves job performance

Professional eye care, and the adaptation of optical aids to individual needs and job demands, has given workers an advantage in various activities. Experiments in the use of special types of occupational lens correction for jobs requiring very close vision have been carried out in England.[29] These glasses include a component, over and above what would be indicated for ordinary lens corrections for these persons, that has the same optical effect as though the work were removed farther from the eyes but without loss of visibility.

---

[29] H. C. Weston and S. Adams, *On the Relief of Eyestrain Among Persons Performing Very Fine Work* (London, H. M. Stationery Office, 1928), and *Further Experiments on the Use of Special Spectacles in Very Fine Processes* (London, H. M. Stationery Office, 1929).

In psychological terms these lenses modify the visual requirements of the job so that they come within the range of adaptability of a larger number of potential operators. These spectacles are worn *only on the job;* they are part of the job equipment or tools. At other times each operator wears whatever spectacles may be desirable for general use, or, if none are needed, no spectacles at all.

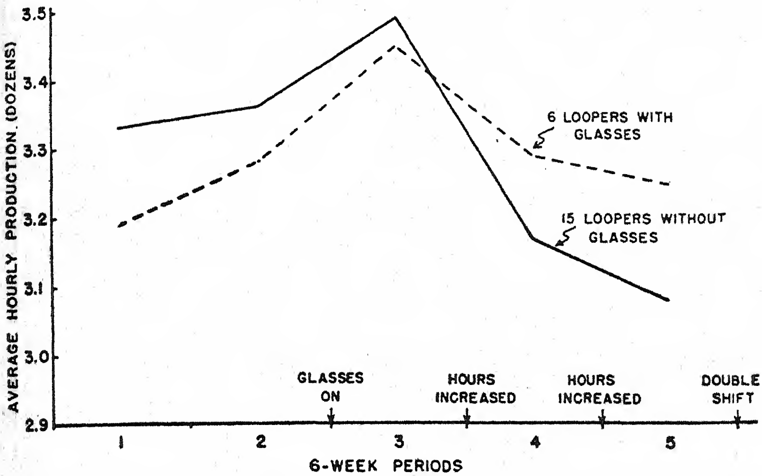


FIG. 85—Effect of occupational eyewear on the production of learning loopers.

These early experiments have been duplicated under the medical supervision of Dr. H. S. Kuhn on one job in hosiery manufacture—the operation of looping. This job requires an average visual working distance of eight inches. Figure 85 shows the production of learning loopers on coarse-gauge hosiery who wore such spectacles, in proportion to the production of a comparable group who did not wear such spectacles. At the outset, the control group was producing more than the experimental group, while both groups were improving. With two increases in hours worked, these learners adjusted their hourly pace downward. At the conclusion of the experiment the group wearing glasses had gained 0.3 dozen pairs production per hour more than the control group

—a gain of 10 per cent in production and piece-rate earnings. This advantage enabled them to reach their peak of production—which was also a higher peak—more rapidly than the group not wearing glasses. Since this one operation of looping accounts for about 20 per cent of the labor cost in the manufacture of circular-knit hosiery, more economical operation through the use of such spectacles would seem to be possible.

The spectacle prescriptions for these experiments varied from one operator to another because of individual differences in visual characteristics, although the job requirements were uniform.[30] Not all of the operators would or could wear such spectacles successfully, but the advantages might become available to a larger proportion of such operators with improved bases for predicting individual lens prescriptions for occupational use. Thirty operators who wore these spectacles successfully showed on a visual retest six months later a pronounced and uniform shift of visual characteristics away from the peculiar conditions that characterize most operators on this job and toward what is generally considered normal; and their subjective reports indicated in most cases less fatigue during work and greater pleasure from activities involving the use of their eyes after work.

Such specialized occupational eyewear cannot be introduced in any industrial situation except by a competent and ingenious ophthalmologist or optometrist who is not only qualified as an eye expert but also is familiar with the visual problems of specific jobs and with typical management attitudes and methods of approach to such problems. Ideally such a project should be carried on in collaboration with an industrial psychologist who can set up adequate criteria as a goal. The first step is to demonstrate a relationship between

[30] The optical addition was varied by Dr. Kuhn for different employees according to the results of an examination. The average optical addition was an amount that changed the focus requirement from 8 inches to 11.5 inches and the convergence requirement from 8 inches to 9.5 inches.

certain visual characteristics and service or success on the job. For example, preliminary studies on other jobs in the manufacture of full-fashioned hosiery indicate that occupational glasses like those for loopers might be generally beneficial for menders but not so beneficial for seamers, for whom, possibly, some other type of special occupational eyewear would be of benefit. The second step in carrying on such a project is to determine how certain visual characteristics may be induced safely and beneficially by lenses. The final step is to evaluate the success of the project.

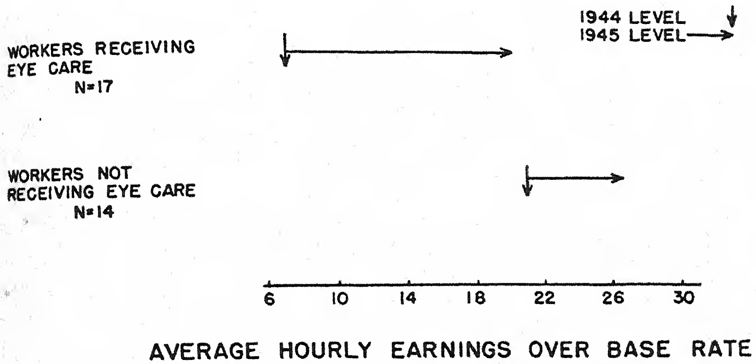


FIG. 86—Increased earnings of hosiery pairers as a result of eye care.

In another hosiery mill all women employed in matching stockings exactly into pairs had been tested on the Ortho-Rater. Minimum visual requirements had been established for this job, and those employees who did not meet the requirements were referred for eye care to the professional eye men of their choice in the community. The results of a follow-up study a year later, reported by Kephart,[31] are shown in Figure 86. Earnings were based on production at piece rates. A change in method enabled all pairers to produce and earn more. Those who had not received eye care had increased

[31] N. C. Kephart, "An Analysis of Professional Eye Care and Industrial Efficiency," *Transactions of the American Academy of Ophthalmology and Otolaryngology*, L(1946), pp. 166–170.

their earnings approximately 5 cents an hour. Those who had received eye care, though lower in earnings at first, had increased their earnings 13 cents an hour—a gain of 8 cents an hour over the other group, or a yearly income increase of nearly $200.

Results from a similar follow-up study of electric solderers employed in joining the small parts of gold spectacle frames, reported by Wirt,[32] are shown in Figure 87. Reconversion from war to peace was accompanied by an average loss of production and earnings of 4 cents an hour for those who had

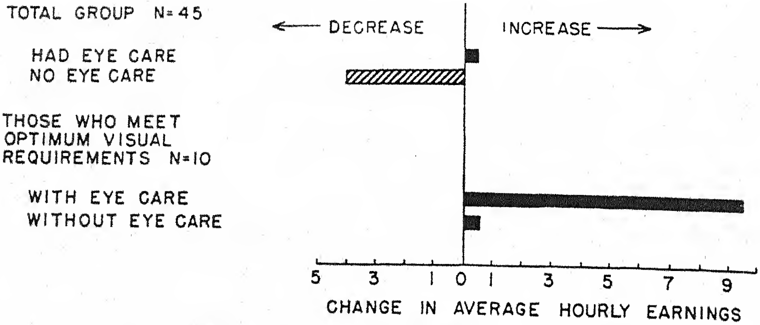


FIG. 87—Effect of eye care on earnings of electric solderers.

received no eye care, but those who had received eye care held their own and actually increased their earnings one-half cent an hour. Among those cases who could qualify on the most severe, or optimum, visual requirements for the job, those without benefit of eye care increased their earnings a fraction of a cent an hour, while those who had benefited by eye care increased theirs by 9½ cents an hour. These studies show that vision can be improved, that it can be adapted to specific job demands, and that eye care tends to increase production and earnings, especially when it enables the employees to meet the visual demands of the job.

---

[32] S. E. Wirt, "Eye Care Increases Earnings of Electric Solderers," *Optometric Weekly*, XXXVII (1946), pp. 1091–1093, 1104.

Visual Maintenance

Proper placement of employees is only one part of a program for taking advantage of visual differences among employees. The other part is proper maintenance of these skills. Good practices in simplification of jobs, adequate lighting, and use of various types of occupational aids will help in the maintenance of adequate visual performance among employees. Proper job lighting should give attention not only to general illumination but also to specific job requirements for illumination. Standard general practices for lighting have been codified by illuminating engineers.[33]

Eye protection

Adequate physical protection of the eyes of workmen also helps in maintaining standards of visual performance on the job. Visual hazards in industry that threaten visual competence of employees include the possibilities of sudden injury (trauma) or of cumulative ill effects like industrial diseases, such as ultraviolet "burns" among welders. Some visual hazards in industry, against which protection is possible, are:

1. Ultra-violet radiations (arc and acetylene welders, welders' helpers).
2. Infra-red radiations (melters, furnacemen, and others working around furnaces and molten metal).
3. Air-borne dust and noxious fumes (sand blasters, colliers, millers, trainmen, truckers, aviators, workers in some chemical operations).
4. Splash and spatter (welders, grinders, handlers of acid and other caustic chemicals, solderers, foundrymen).
5. Severe impact and explosion (riveters, chippers, chemists, drillers, powder men).

Each of these hazards requires a different type of eye

---

[33] Committee on Lighting Practice, Illuminating Engineering Society, *Recommended Practice of Industrial Lighting*, The Society, New York City.

protection, including masks, optical filters, and lenses highly resistant to breakage and shattering. Minimum standards for such protective devices have been formulated by government agencies.[34] Manufacturers have made progress toward adapting the design of such devices to specific job requirements of vision and visibility, to individual differences in facial size and contour, and to various combinations of hazards. The economic value of such devices has been demonstrated repeatedly.[35]

Although such measures for the maintenance of visual proficiency are taken, it is still certain that changes in vision will take place in a majority of employees with increasing age and job experience. Some of these changes, such as the loss of near vision with age, have long been accepted as inevitable. Yet even this change, universal and serious as it is for individuals who must do any reading or close scrutiny on their jobs, is often not recognized and corrected until the efficiency of the employee has been definitely reduced. Perhaps one reason for this is that the loss of near vision with age does not occur suddenly like a broken arm or a sprained ankle, but develops so gradually that an employee is often unaware of his increasing handicap until long after his work has been seriously affected. Indeed, paradoxical as it may sound, no one can tell from the way things look to him whether his vision is normal or below normal. Just as many persons go through life without knowing that they are color-blind—unless they have been tested and told the result—so also many persons with seriously deficient visual performance in other respects are unaware of their real condition unless it is demonstrated to them in terms of objective test results. These facts make

---

[34] U. S. Department of Commerce, National Bureau of Standards, *American Standard Safety Code for the Protection of Heads, Eyes, and Respiratory Organs*, National Bureau of Standards Handbook H24 (U. S. Government Printing Office, 1938).

[35] H. P. Davidson, "Two Years of Industrial Ophthalmology at the Pullman Car Works," *Journal of Industrial Hygiene*, VIII (1925), pp. 247–253.

it necessary to check periodically the vision of present employees.

Job simplification

An important part of an industrial program of vision is the recognition and simplification of visual operations in the plant and the improvement of visual working conditions. Visual factors in job operations can often be eliminated. Where they cannot be eliminated altogether, sometimes simpler visual functions can be substituted for more complex ones. Jigs and fixtures for positioning tools and materials help to reduce visual operations to a minimum. Another factor in the simplicity of visual operations is the area over which they are performed. Several studies in the Purdue Time and Motion Study Laboratories have shown the costliness of eye movements in relation to time expended and accuracy of work achieved[36]. Some visual cues can be simplified so that they can be observed with peripheral or "side vision" without turning the eyes. Figure 88 shows the spread of visual operations on two repetitive jobs. In general, the visual operations on any job should be organized into a restricted area at a convenient and fairly uniform distance from the eyes. The visual working distances can best be measured from the plane in front of the eyes where spectacles are usually worn (the lower forehead or base of the nose approximates this position) to the various points of visual attention on the job. The angular deviation of these "lines of vision" from horizontal is also important in connection with the use of spectacles, especially bifocals or special occupational glasses. Such information may be helpful in determining visual requirements for various jobs, and should be incorporated in job analyses and descriptions for use of eye doctors in prescribing lens for specific jobs.

---

[36] J. Tiffin and H. B. Rogers, "The Selection and Training of Inspectors," *Personnel*, XVIII (1941), pp. 3–20.

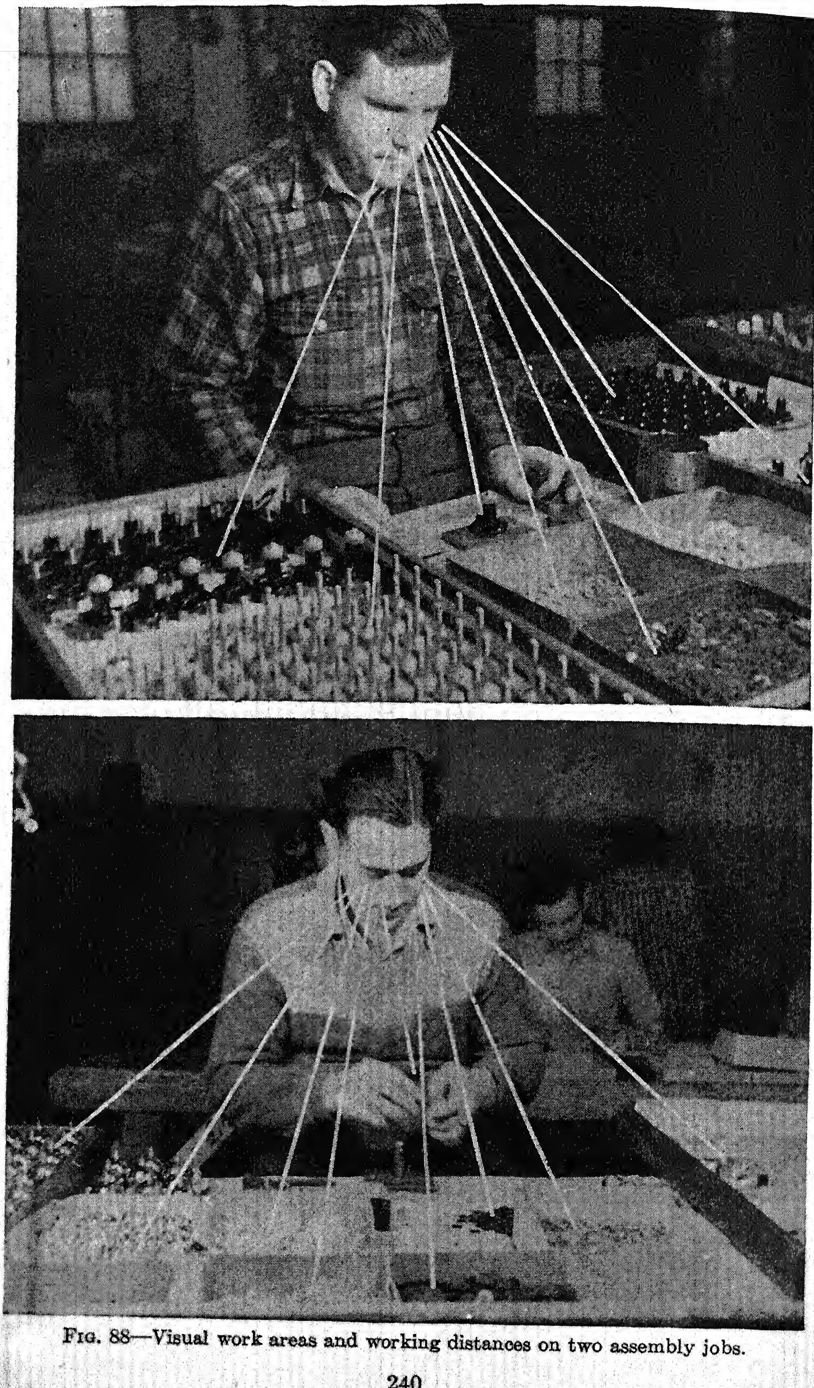


Fig. 88—Visual work areas and working distances on two assembly jobs.

Optical aids

While job simplification and adequate lighting may reduce the visual requirements on some jobs, other visual requirements can be modified by the use of optical devices that permit a larger proportion of available employees to perform a job satisfactorily. Special types of microscopes, magnifiers, and micro-photographic apparatus permit visual inspection and scaling of very small surfaces and pieces and analysis of material structure. Small machined parts can be magnified in profile for direct comparison with original large-scale drawings by means of the contour projector, shown in Figure 89. Polariscopes reveal stresses and strains in transparent models of castings. These are only a few of the many types of optical instruments, developed for specialized work in industry, that, like gauges, meters, and dials, bring important manufacturing operations within the range of visual observation of average employees.

Without colored glass (optical filters) some operations, such as welding and melting, would be practically impossible. The use of clear, hardened protective glass enables employees on other jobs, such as chipping, sanding, planing, drilling, and grinding, to face their work without flinching or averting their eyes, thus making full use of their visual capacity. The importance of spectacles for the compensation of individual anomalies of vision is shown on the preceding pages. Because of the differences in visual performance for different distances, mentioned on page 239, such spectacles would seem to be most appropriate for industrial purposes when they are designed to meet not only the individual, personal needs but also the specific visual requirements of a particular job at some particular work distance or range of distances.

Periodic retests of vision

Employees in certain critical positions in industry can usually be required to maintain good vision. Operators of mobile equipment have responsibilities for the safety of other

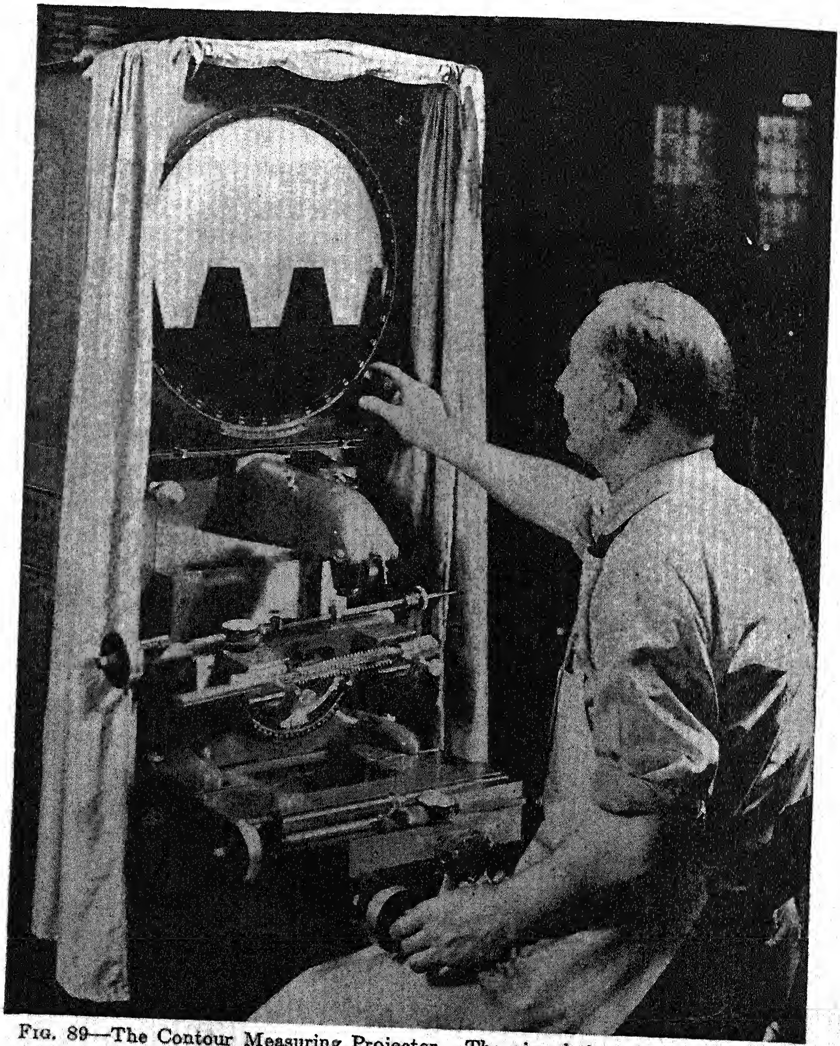


FIG. 89—The Contour Measuring Projector. The view before the operator is an enlargement of the teeth on a machine tool.

employees. Supervisors have responsibilities for correctly interpreting instructions to workmen and maintaining both quality and quantity of production. Maintenance and repair men have responsibilities for keeping other men from becom-

ing idle on account of breakdown. Inspectors have far-reaching responsibilities for quality of product. These and other employees whose functions directly affect other groups of workmen or the general welfare of the plant can and should be made to feel their responsibilities with respect to maintaining their own physical functions at a high level of efficiency. Employees in some industries, such as the common carriers and other public utilities, whose operations are subject to public observation, frequently show great pride in the physical qualifications that they maintain and demonstrate in repeated examinations.[37] Visual requirements for a license to drive a car in many states have given a great impetus to such attitudes.

Periodic retests are usually administered in the same place and follow the same routine as the pre-employment examination. At designated periods, and upon due notification, an employee presents himself for re-examination. At that time he is usually interested and ready to discuss his personal health and efficiency and to take any steps that are advisable to maintain them. This feature of a visual program can well be inaugurated with supervisors (to set the precedent), extended to cranemen, truckers, and others in the field of transport within or outside a plant, to inspectors, to gang leaders, to men on hazardous jobs, and to any other group that has an easily recognizable visual responsibility. The interval between tests need not be uniform for all groups. Indeed, it should be determined separately for each group in proportion to the importance of vision in that group.

These retests cannot be done all at once, but they can be distributed over a long period and thus made a small part of routine, rather than special, activities. This usually involves a separate file for each group of employees arranged in calendar order of their next examinations. Sometimes, however, greater economy in such testing operations can be achieved

---

[37] Hart E. Fisher, "What Periodical Examinations Can Accomplish," *National Safety News*, XXXII (Dec., 1935), pp. 33–34.

by consolidating them into a limited period at a specific time.

More general vision testing among employees in a plant or a department may be advisable from time to time in an attempt to combat the losses that occur with age and to raise average levels of visual proficiency in the plant. Such gains are effected in proportion to the number of employees whose vision can be improved somewhat rather than to the amount of improvement that can be achieved in the vision of a few employees. It is important and economical to retain employees with experience on the job, to "salvage" as many as possible of the employees whose visual proficiency has deteriorated, and to find other jobs for those who have become visually incapacitated for their present jobs.

The immediate objective of a visual survey is to identify employees whose visual performance is low and to stimulate them to obtain professional attention. Since such professional treatment requires both time and money, it becomes necessary to select those employees who are most in need or who would be most likely to be benefited by such services. Some of these employees can be selected on the basis of their own complaints or their progressive loss of proficiency on a job requiring visual ability; but too often a serious visual deficiency will long be undetected by these means.

Maintenance of occupational eyewear

Any glass before an employee's eyes should not only be the best glass for his needs but should be maintained in the correct position for maximum value. In a survey of employees in a steel mill only one-half of one per cent of employees showed a serious tendency for one eye to point slightly higher than the other; but when this condition occurs it is often the forerunner to double vision during periods of fatigue. Among the employees who were not wearing glasses, .42 per cent showed this tendency, but among those wearing glasses, .78 per cent showed it. This condition can be induced by glasses that are not maintained level and in the correct position. Apparently

this fact accounts for the greater percentage of this tendency among employees wearing glasses.

Employees who have obtained their glasses on prescription from a reputable eye man have in most instances thereby established contact with adequate services in adjusting and maintaining their glasses. But many employees do not have such contacts or do not make use of them frequently enough. A few industries have hired men permanently to supervise employee eyewear. Such a supervisor selects the right kinds of safety eyewear and fits them to an employee's face, makes adjustments and repairs on all types of eyewear, makes up special devices for special purposes, cleans and sterilizes safety eyewear, and instructs employees in the use and maintenance of their eyewear. He does not make professional examinations or fill prescriptions. Such a man can pay his own way in a plant by the savings he effects from salvage of damaged goggles. The total economic saving from his work is likely to be many times the cost.

Employee education in vision

It should be evident from the preceding discussion that a visual program for employees now in service is a phase of personnel maintenance and therefore involves problems of employee training or education. Employees must be motivated to develop habits of personal maintenance, to expect to meet certain standards of visual proficiency, and to discover for themselves a source of professional eye care that can keep their visual proficiency at a maximum. The visual testing itself and the experience that it involves are important elements in this training. The effect of survey tests as a frankly educational device in improvement of job performance, apart from any other corrective measure, has been demonstrated,[38] and many reports of improvement in production, safety, and

[38] E. D. Fletcher, "Effect of Special Tests on Driving Ability," *Preliminary Report on Special Tests* (State of California, Department of Motor Vehicles, Division of Drivers Licenses, 1939).

employee morale as a result of visual surveys can well be attributed in part at least to this factor. It is reasonable to expect, however, that any real improvement in job performance due to improvement in visual proficiency will be reflected in improved performance on vision tests.

A special problem in vision education is employees' attitudes toward vision tests and visual standards. Employees often have only a few vague notions and many misconceptions as to the scope of visual functions, the nature of visual deficiencies, and the significance of optical aids. A concept widely in use at a former time, but still heard among older persons, is "weak eyes." Glasses were for "weak eyes," and those who wore glasses had "weak eyes." This term no longer has any true or specific reference among the concepts of professional men and has been replaced largely in popular thinking by other terms. A current term, under which are included most problems of vision, is "visual defect"—professional ocular services and optical aids "correct" visual defects. This seems an unfortunate concept because of the universal resistance to admission of personal defect. A person who solicits ocular services and accepts optical aid for his eyes must go through the mental operation of admitting and accepting a personal defect—unless he can find some other rationalization for his conduct. This attitude has been further promoted by visual surveys and visual standards in employment, automobile driving, and other areas of activity, since those who are selected for ocular attention by these methods show substandard performance, and the corollary implication is that ocular attention is primarily if not exclusively for those with substandard or "defective" vision. The result of this attitude is that many workmen deny themselves ocular advantages because of pride. In an attempt to overcome this psychological handicap to visual improvement, concepts of "visual efficiency" and "visual skills" have been popularized in recent years.

An industrialist responsible for employee education should

be able to help an employee understand his visual problems and to appreciate the purposes of industrial vision testing so as to elicit maximum co-operation toward improving his employability. From an industrial standpoint, visual conditions are considered functionally as elements of aptitude or adaptability. Whether the limits in visual adaptability are functional, structural, or pathological is not the concern of the psychologist—nor, for that matter, of anyone except the professional eye man. The industrial problem is to discover such limitations *insofar as they are related to industrial performance and insofar as they can be determined reliably by simple, non-professional tests,* such as those mentioned on page 185. These tests may be thought of as sampling different aspects of visual adaptability. Limitations in such functions are to the professional eye man symptoms of some basic dislocation of visual responses. To the industrialist they represent individual differences in visual aptitude or adaptability for specific job requirements. Visual aptitude, in the sense of adaptability, is different from aptitude as ordinarily defined, for visual adaptability can often be modified in many respects by means of spectacles, orthoptic training, magnifiers, microscopes, and other optical aids. Eyeglasses mediate between the visual demands of the job and the visual capacity of the worker. The end result should be better efficiency in activities of major importance.

The employee can be encouraged to anticipate functional improvement in his visual characteristics and to look upon glasses or other methods of eye care as a means to this end. If he wishes to verify the result, after a period of re-adaptation, in terms of the same industrial tests that he originally failed to pass, he should have an opportunity to do so.

Another problem of education in visual hygiene is the attitude of the employee toward using occupational eyewear of various types. In some instances, protective or corrective spectacles may be actually uncomfortable or otherwise not adapted to an employee's individual needs and conditions.

But a more general unreadiness to accept optical aids of this sort can be traced in part to the popular misconceptions previously mentioned and in part to employment policies that have sometimes been followed. In many industries, an applicant wearing glasses has been considered ineligible to undertake a physically active job. Surreptitious use of spectacles has been reported[39] among workers in the building trades. On jobs where ordinary spectacles are liable to breakage their use is certainly unsafe, not only because of the hazard to the wearer's eyes but also because an employee who has broken his customary glasses is visually unadapted to his job and may therefore endanger himself and others. With the development of combination corrective and protective lenses, or "prescription goggles," the hazard of glasses is not only eliminated but eyes are actually safer behind such lenses than they are without them. Corrective spectacles, whether they have safety lenses or not, play an important part in progressive industry today.

"Getting glasses" is not necessarily the answer to the visual needs of employees. Neither is it the right recommendation to make to employees who show substandard performance on industrial tests. Some basic causes of visual deficiency must be attacked by other means than eyeglasses, and many functional deficiencies cannot be remedied by spectacles alone. The right recommendation is to get professional eye care: a competent examination, treatment, properly fitted glasses if and as prescribed, and any necessary after care.

---

[39] Francis E. Carroll, "Defective Eyesight in the Building Trades," *The Bausch and Lomb Magazine*, XVI, No. 3 (1940).

8

Training of Industrial Employees

EMPLOYEES must be systematically trained if they are to do their jobs well. No matter how carefully men have been selected, or how much aptitude they may have for their jobs, systematic training is essential if they are to reach a satisfactory level of job performance. Formerly it was standard practice to have training done by the foreman or supervisor in charge of the work. Inexperienced men were sent to the supervisor (or hired by him) and it was his job to see that the new men were given the necessary training. But just as modern scientific management has found that experts should supplement the work of the foreman in making job analyses and setting rates, so also management is finding that the use of experts in training is advantageous as a supplement to the work of the foreman. This does not mean that outside experts need necessarily *do* the training. Their function, rather, is to set up the machinery for the training to be done by properly qualified persons, to supervise the construction or writing of necessary manuals or other materials, and to keep in close touch with the accomplishment of the trainees during the instructional period. Through this procedure training costs of various departments and various jobs may be compared, and the relative efficiency of different training methods may be evaluated in an objective (and often financial) manner.

The Measurement of Training

A comparison of training methods, or the evaluation of any single method, requires some means of measurement.

The value of precise means of measurement of a type thoroughly accepted in other branches of management is now being recognized as of equal value in industrial training. There are several standards that may be used in measuring the effectiveness of training.

Production

Training is usually training for production, and therefore the success (or failure) of the training is usually indicated by the production achieved by the trainees at various stages of their training.

In many industrial jobs the work has been time-studied, rates based on these studies have been set, and output or production of employees is indicated by production or earnings. A comparison of production from one time to another reveals immediately how effective the intervening training has been and whether new or different methods should be tried out. A typical example of this method has been reported by Blankenship and Taylor[1] for the operations of trimming, covering, and hemming among textile workers. The learning curves for these three operations, as published by Blankenship and Taylor, are reproduced in Figure 90. In order that these curves may be comparable from one operation to another, the production in each case has been plotted in terms of Bedaux units. These units consider the actual motions involved in each operation and equate these so that a given number of Bedaux units means a given amount of production, regardless of the operation involved. Thus, thirty Bedaux units of production in trimming may be considered equal to thirty Bedaux units in covering or hemming.

It will be noted that for each operation employees with less than five weeks of experience produced from thirty-one

---

[1] A. B. Blankenship and H. R. Taylor, "Production of Vocational Proficiency in Three Machine Operations," *Journal of Applied Psychology*, XXII (1938), pp. 518–526.

to forty-one Bedaux units. With increasing experience, the output in each case rose until, for trimming and hemming, it reached sixty-nine units at the end of thirty-five weeks; while in the case of covering, only sixty-two units were reached in this length of time. It is apparent that no serious discrepancy in speed of learning exists among the three

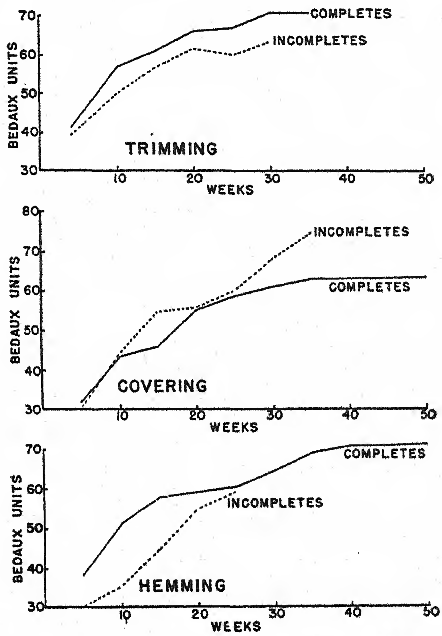


Fig. 90—Learning curves for three operations in a textile mill.

operations studied. It will also be apparent, however, that if a serious discrepancy did exist, that is, if employees on any one of these jobs tended to lag consistently behind employees on the other jobs in speed of learning, an exact quantitative measurement of the amount of the lag would be indicated immediately by learning curves of this type. Such curves, therefore, serve the very useful purpose of providing a means of evaluating the successfulness of an operator-training pro-

gram and spotting decisively those operations in which training is inadequate, either in quantity or quality.

The question may be raised as to whether such curves as those in Figure 90 indicate the effect of *training* or the effect of *experience* on the job. The answer depends on the type of activity that has taken place during the period studied. Obviously, if no formal training has been given and each

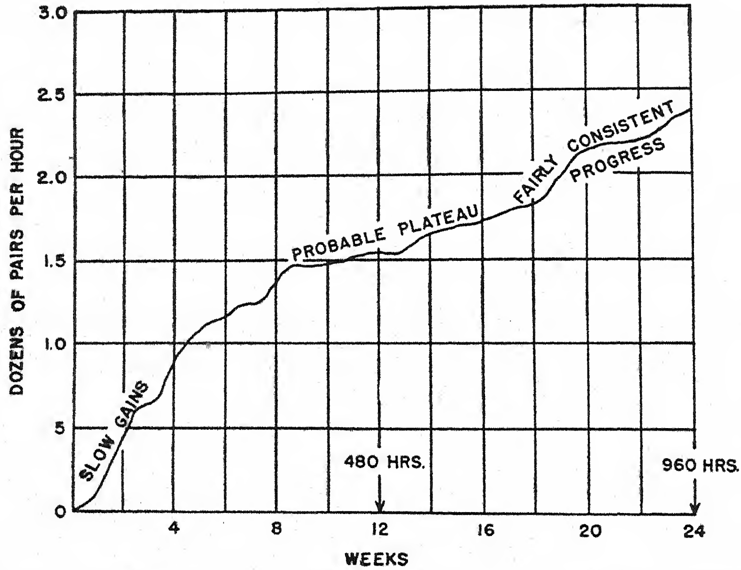


FIG. 91—Learning curve of hosiery loopers. (By courtesy of J. D. Boyd, Personnel Director, Melrose Hosiery, High Point, N. Carolina.)

operator has been left free to pick up the "tricks of the trade" in whatever way he or she can, the curves would indicate only the improvement due to experience. On the other hand, if formal training has been given systematically over a period of weeks, the curves will indicate the amount of improvement in job performance occurring as a result of the training and experience. The curves simply show the change in job performance that takes place. Whether this change is due to training or experience or to a combination of these factors can

be decided only in the light of a complete knowledge of the conditions under which the operators have been working. But when the curves for a given operation are available, together with a knowledge of the working conditions, it is possible to evaluate in objective and accurate terms the effectiveness of the method that has been used in training new operators.

A similar curve showing the relationship between production and experience on the job of looping in a hosiery mill has been reproduced on page 18 in connection with the discussion on individual differences among operators at different stages of the learning process. Another curve of learning for hosiery looping from another plant is reproduced in Figure 91, a copy of which is given to all new employees on this particular job in this plant. Although these curves have not been graphed in terms of Bedaux units, the general principles involved in their construction are the same as those just discussed.

Production time required to do the job

It is sometimes more convenient to measure the effectiveness of training in terms of the reduction in time required for the job as the training progresses.

An excellent example of the application of this method to the evaluation of a job-training program has been published by Greenly.[2] Figure 92, reproduced from Greenly's article, shows the reduction in time required to change the knives on a flying shears following a program of job training dealing with this operation. The graph shows that the average time required for the operation for the year preceding the installation of the training program was twenty-nine minutes. Immediately after the training was undertaken a drop in the time required for the operation is apparent. This drop continued for three successive months, at the end of which

---

[2] R. J. Greenly, "Job Training," *National Association of Manufacturers Labor Relations Bulletin, No. 35* (1941), pp. 5–8.

time the operation required an average time of only eighteen minutes. From a graph of this type it is a simple matter to compute the dollars-and-cents savings effected by such a job-training program. In this instance the twelve minutes

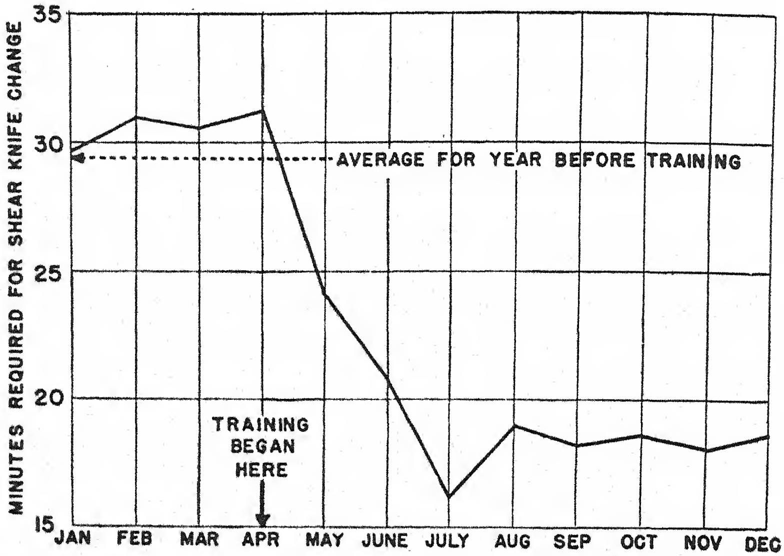


FIG. 92—Effect of a job training program on the time required to change knives in a flying shears.

saved times the minutes labor cost times the average number of changes resulted in a saving of $1740 per month or $20,880 per year.

Number of operators meeting production standard

Where production standards have been set, a measure of the effectiveness of training is the number (or percentage) of trainees who meet the standard after a given period of training. If a new training program is initiated, its value may be checked by a comparison of the percentage of operators who meet production standards after a given period of time with the corresponding percentage of operators who had met the standard after the same period of training under the

previous program. This method of measuring the effectiveness of training is illustrated in Figure 93, published by Lawshe.[3] In this figure the bar graphs indicate the percentages of operators meeting a given production quota after four and after eight weeks of training. The bars show the results for a group of operators in a new training program and for a group not in the program. It is apparent that operators

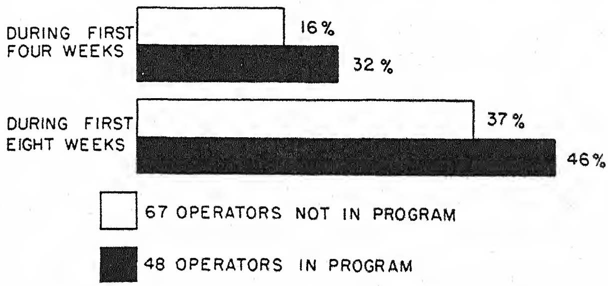


FIG. 93—Effect of an "induction" program upon per cent of employees engaged in a record press operation who met production standards after four and eight weeks, respectively. (Figures 93 to 98, inclusive, are reproduced by courtesy of *Factory Management and Maintenance* from C. H. Lawshe, Jr., "Eight Ways to Check the Value of a Training Program," *Factory Management and Maintenance*, May, 1945, pp. 117–120.)

in the program excel those who are not, both at the end of four weeks and at the end of eight weeks.

Increase in learning rate

Comparison of the learning curves for an operation under two sets of learning conditions provides another means of evaluating the relative effectiveness of two training programs. Such a comparison is shown in Figure 94 from Lawshe.[4] In Figure 94 are plotted the earnings of apprentices at the end of eight, successive six-month periods under a piece-work method of payment, in comparison with earnings at corresponding periods of a group of apprentices paid the customary contract

[3] C. H. Lawshe, Jr., "Eight Ways to Check the Value of a Training Program," *Factory Management and Maintenance*, May, 1945, pp. 117–120.
[4] *Ibid.*

rates. It will be noted that under the piece-work plan, the apprentices are earning the same amount at the end of thirty months (five six-month periods) as the apprentices under customary contract rates earn after forty-two months.

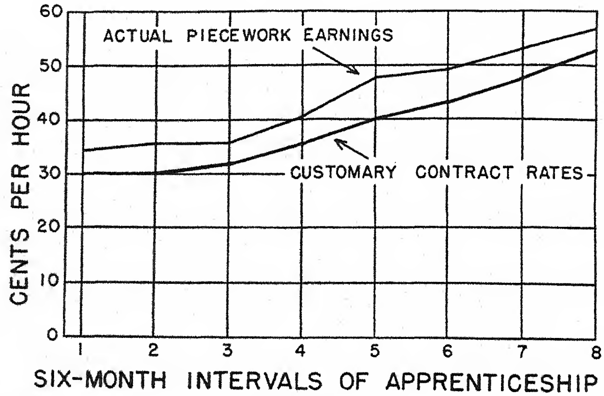


FIG. 94—Production of apprentices paid on an incentive plan in comparison with those paid customary contract rates.

Decrease in breakage or use of consumable supplies

If two training methods result in a significant difference in use of tools, parts, and other consumable supplies, the method involving the smaller use of supplies is ordinarily the better one. This quantitative method of measuring the effectiveness of training is illustrated in Figure 95.[5] The operators whose work is charted were on the job of cutting small tungsten discs on a machine using a rotary abrasive cutting wheel. One of the major costs of the operation was the use and replacement of these wheels. In Figure 95 the horizontal dotted lines show the per cent of wheels broken by previous employees, who were not in the training program, at the end of eight weeks, twenty weeks, and thirty-six weeks, respectively. It is apparent from the chart that in each case the

[5] L. G. Lindahl, "Movement Analysis as an Industrial Training Method," *Journal of Applied Psychology*, XXIX (1945), pp. 420–436.

trainees broke fewer abrasive wheels than did previous operators at the end of equal periods on the job.

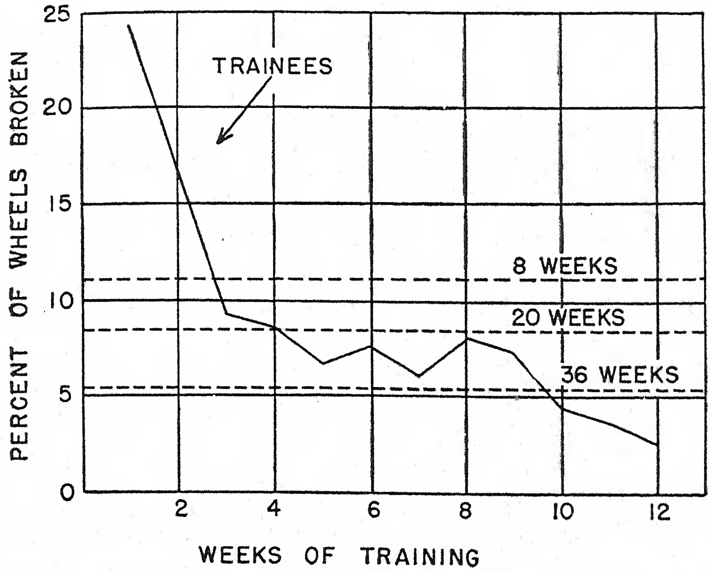


FIG. 95—Decrease in abrasive wheel usage of trainee group in comparison with present employees of varying amounts of experience.

Reduction in number of accidents

In many plants and on many operations, accident prevention is one of the basic matters covered by a training program. In such situations, a means of measuring the

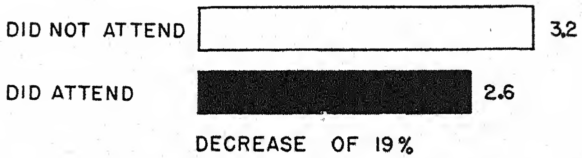


FIG. 96—Comparison of accident experience during first 30 days of employment of employees attending a vestibule training program and employees broken in "on the floor."

effectiveness of training may be found in a comparison of accident experience data before and after the beginning of a new training program. This type of measurement is illustrated

in Figure 96 from Lawshe.[6] In this chart are plotted the percentage of employees having one or more accidents during their first 30 days on the job. The top bar is based on the record of the last 70 employees broken in "on the floor," while the lower bar is based on that of the first 70 employees to go through a new vestibule training school.

Reduction in absenteeism

A similar measurement of the effectiveness of training may often be obtained from records of absenteeism. This method is illustrated in Figure 97, also from Lawshe.[7] During the

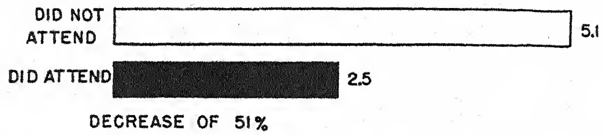


FIG. 97—Comparison of absenteeism during first 30 days of employment of employees attending a vestibule school and employees broken in "on the floor."

first 30 days of work for each group, there was a total of 152 days of absence among those who were trained on the floor, in contrast with a total of only 74 days of absence among those who had the vestibule training.

Reduction of labor turnover

The expense of training new men is usually a significant part of production costs. Other things being equal, a training

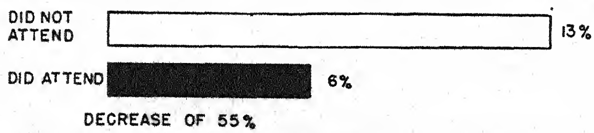


FIG. 98—Comparison of terminations before 30 days of employment of employees attending vestibule school and employees broken in "on the floor."

program that keeps labor turnover at a minimum should be sought. A further measurement of the effectiveness of the

[6] Lawshe, *op cit.*
[7] Lawshe, *op. cit.*

vestibule school referred to above was obtained from data on turnover before and after the school was inaugurated. These results are graphed in Figure 98, from Lawshe.[8]

Company policy and technical information

To measure the effect of training in terms of immediate monetary savings, using one or more of the indices described above, is not always possible. When that is the case it does not mean, however, that measurement is impossible. On the contrary, measurement may often be made of the effectiveness of training in such matters as company policy, knowledge of job, technical information, and so on, by means of specially constructed tests. For example, many industries devote a considerable amount of attention to supervisory training conferences. Such conferences frequently emphasize the importance of company policy and devote a considerable amount of time to instructing supervisors in ways of handling numerous kinds of problems, the solution of which requires a knowledge of company policy. It is not necessary to *assume* that such conferences implant the necessary or desired information in the mind of the supervisor, nor is it necessary to rely upon the supervisors' answers to such questions as, "Have these conferences improved your ability to handle situations?" in order to find out whether the conferences have actually had the desired effect. A more satisfactory method of evaluating such conferences is to resort to the general principle of measurement. Such measurement can readily be made by means of specially constructed true-false or multiple-choice questions that cover thoroughly the subject matter of the conferences and that determine in quantitative terms just what the supervisors know and do not know about the subject matter before and after the conferences have been held. Typical examples of the kind of questions

---

[8] Lawshe, *op. cit.*

that may be used for this purpose have been published by Mapel.[9]

If an employee is called out for work at 8:00 P.M. because of a vacancy in a position regularly filled by another employee, and he works until 12:00 P.M., he should be paid for:

1. 4 hours work.
2. 5 hours work.
3. 6 hours work.
4. 8 hours work.

In department "A" there are three shears; John Jones is first helper on No. 1 shear, who, due to sickness of the shearman, has acquired two weeks experience as a shearman. Several months later a vacancy exists in the shearman's position. First chance at the vacancy should be given to:

1. John Jones.
2. The first helper with the longest departmental service.
3. First helper with the greatest plant service.
4. The first helper with the longest service where combined factors are relatively equal.

Questions such as these result in a *measurement*—not an estimate or a guess—that reveals whether the men who have been trained actually have achieved an increased understanding of company policy and the union contract.

Questions of this general form may be used in still other areas of training to determine the effectiveness of instruction. Many industries have in-service training that covers technical information or content peculiar to their individual industries. The trade of pipe fitting, for example, though many of its details are common in various industries, often involves certain particular information of a technical nature that is required of the pipe fitters in a given industrial plant. If training covering this specific technical information has been given, it is only reasonable for management to ask whether

[9] E. B. Mapel, "Stimulating Employee Self-Improvement," *Personnel Journal*, XIX (1941), pp. 316–324.

the training has been effective. This can be determined by such questions as the following, taken from a typical test used in a large steel mill:

The correct pressure to use when unloading an acid tank car is (1) 27 (2) 36 (3) 64 (4) 90 pounds. ( )

The correct pressure to use when unloading an oil tank car is (1) 27 (2) 36 (3) 64 (4) 90 pounds. ( )

Oil and grease must be kept out of a pipe used for explosive materials because (1) contaminated gases do not burn so well, (2) an explosion will result, (3) it will cause clogging of the burners, (4) it makes repair work more difficult. ( )

City water lines should not be connected to mill water lines because (1) this would contaminate the city water, (2) city water is more expensive and should be used for drinking only, (3) the chlorine in the city water would damage machinery and steel product, (4) mill water is under high pressure and this would burst the city water lines. ( )

The foregoing illustrations will suffice to show that whatever may be the subject matter of industrial training it is possible to devise some sort of quantitative method to measure its effect. The measurement may involve any one or several of the methods discussed, and in some instances it is, indeed, desirable to utilize several to measure the effectiveness of training. Management need no longer accept a training program on faith or because of someone's flowery statements about the presumed value of the program. The application of psychological measurement makes it possible to evaluate training objectively. Without such measurement, there is no way of knowing whether the training is worth what it costs, or, indeed, whether it is worth anything at all.

Types of Training

Job training of present employees

Modern industry is in a process of continuous technological change. No matter how well an employee may be able to do his work today, tomorrow he may find his job so changed

that additional training is needed if he is to remain efficient. Continuous job training is a necessary corollary of continuous technological change. Job training to keep present employees efficient on their jobs was formerly conducted by supervisors. The ever-increasing complexity of industrial processes has caused the supervisor to become less and less able to cope with the technical phases of this training. In 1916, nearly any handy man with a few tools could do a reasonably satisfactory job of repairing his own automobile, even to the point of overhauling the motor. But the modern automobile has become so complex that a mechanic without specific training in the particular adjustments and tolerances recommended by the factory is quite unlikely to accomplish anything like a satisfactory job of overhauling. This problem has been met by supplying factory-trained mechanics and specially prepared manuals, and by offering refresher courses to cover new developments.

Modern industry finds itself in a position similar to that of the automobile service station. No longer can any mechanically minded man be expected to perform satisfactorily in the operation of producing units. And no longer is the foreman able to analyze for himself the intricacies of the machines and to train new operators without outside help. This situation has resulted in the development of training as a new branch of management. The job trainer is essentially a man who has been "factory trained" in the maintenance or operation, or both, of certain producing units. His job is to relay this information by specific training to the employees who are to operate these units. Ordinarily the job trainer is not responsible for production and often he is responsible only indirectly to the foreman or supervisor who is in direct charge of production. In one sense, he is an "outsider," just as the industrial engineer, the safety expert, and the personnel manager are "outsiders." But this does not mean that he should have little contact with the supervisor. He should, indeed, make every effort to develop a co-operative

working arrangement with the supervisors of those departments in which he is doing training. He is not ordinarily allocated for any long period of time to a single department. His job is to work as frequently and for as long periods as are necessary in various departments as an expert who brings technical information and skill to the trainees.

The effectiveness of a well-organized program of job training may be inferred from an inspection of Figure 92, which has been discussed previously. In this instance the improvement in job performance was directly traceable to the activities of job trainers who had been assigned to these specific jobs.

The sources of information for the job trainer himself are varied. In some instances it may be desirable for him to spend a period of time in the factory in which the machine itself is manufactured. By so doing he often learns many things about the operation of the machine that could be learned only with great difficulty, if at all, without this factory experience. It is often desirable for the job trainer to work in close co-operation with the industrial engineering department. By motion-study analysis it is often possible to change the work pattern so that the actual number of motions or complexity of motions in performing the job is considerably reduced. In nearly every instance it is advisable to have the trainer work as a *trainer* while he is on this job and not partly as a trainer and partly as a producer. The illustration cited in Chapter 9 on the training of inspectors typifies the difficulties often experienced when training is offered by an employee who also is responsible for certain production. Of course it is often impossible to allocate any man or men permanently to the job of training. When the training has been completed, the job trainer often returns to his former activities, either as a producer or a supervisor. But usually it is advisable for him to function as a trainer when he is doing training and as an operator or supervisor when he is not doing training, and not attempt to do both jobs at the same time.

The value of job training is so great that just prior to and during World War II one division of the Office of Production Management was maintained specifically to assist industry in setting up and administering such training. This organization was known as the Training Within Industry (TWI) Division of the Advisory Commission to the Council on National Defense. It arranged for the publication and distribution of numerous materials, typical of which is the Manual on Job Instruction[10] of the Western Electric Company.

Job training of new employees

A great deal has been said in the early part of this book concerning the magnitude and significance of individual differences in employee productivity. It was emphasized that these differences are often due to individual differences in employee capacity for the job, which is measurable by appropriate tests at the time of hiring or placement. But it should also be recognized that individual differences in employee productivity are often due in large part to differences in the training that employees have been given. Seashore[11] has pointed out that at least three basic factors may cause individual differences in human abilities. These factors are: differences in biological bases of performance, differences in the amount of previous training, and differences in the particular work methods that are adopted by each individual during the learning period. If several new employees are started simultaneously upon a new job that is somewhat complex in nature, and if these employees are not given identical and standardized instruction in the operations of that job, it is quite unlikely that all will be equally skillful in adopting the most satisfactory work methods. It is exactly as if a number of persons start simultaneously, without

---

[10] *Job Instruction: A Manual for Shop Supervisors and Instructors* (Western Electric Company, 1940).

[11] R. H. Seashore, "Work Methods: An Often Neglected Factor Underlying Individual Differences," *Psychological Review*, XLVI (1939), pp. 123–141.

instruction, to learn the operation of a typewriter. Some, if not most, persons in this situation will adopt the so-called "hunt and peck" system, because this system is easily understood even without instruction and because it results almost at once in some evidence of accomplishment. A few, perhaps, who take the trouble to study the nature of the problem and thus become familiar with the fact that the touch system exists, and who are willing to sacrifice immediate evidence of accomplishment in favor of a learning procedure that eventually will result in considerably greater accomplishment, will not attempt the hunt-and-peck system at all but will start immediately to learn the touch system. After a period of time, say three months, those who learn the touch system will be many times more able to operate a typewriter in terms of both speed and accuracy than those who followed the more immediately apparent method of "hunting and pecking." With full recognition of the greater efficiency of the touch system, every commercial school supervises the training of students in typing so that all are taught the more desirable method. To facilitate the instruction, use is made of such devices as the blank keyboard to insure adherence to the preferred method and to avoid any relapse, even temporarily, to an undesirable method.

In many ways, performance on an industrial job is similar to typing. If employees are placed upon a new job without formal instruction in the operation of the job, each will adopt a method of work that is determined by (1) his insight into what is right and what is wrong, (2) his willingness to forego immediate accomplishment for greater later accomplishment, and (3) what he is able to observe of the activities of other employees on the same job. Management should not expect a new employee to have very much of the first factor mentioned above, that is, insight into improved work methods. Such insight comes only as a result of extended study in the field of motion analysis, the use of special equipment for performing motion analyses, and a considerable

amount of specific study of each job. Neither should management expect employees to have very much of the second factor mentioned above. Each employee when placed on a new job assumes, if he is not instructed otherwise, that he is expected to obtain some production immediately upon that job. It would often be unwise, if not fatal, for him to attempt to teach himself a method that would result in no production for a period of four or five weeks, even if it were possible for him to do so. Management, on the other hand, should certainly not expect employees to have the time, training, or inclination to carry on a prolonged study of the job before attempting to achieve some production. For these reasons new employees will, in the vast majority of cases, adopt an inferior method of job performance if they are not subjected to a formal program of instruction. They will use work methods, to use Seashore's terminology, that will become fixed in the form of work habits that result in generally low production and that later must be unlearned before proper methods can be instituted.

This situation can largely be offset by a program of job training organized so that every new employee is taught from the outset the proper way to do his job. The job trainer must be a man who is thoroughly acquainted with motion study and who either has conducted motion studies of the job or has available such motion studies from the industrial engineering department. He must be a man who is able to teach the proper method to the new employees. The ability to teach presupposes something more than personal skill which he himself may possess. He must be both able and willing to explain the approved methods, be patient while methods are being understood and followed, and be willing to devote considerable energy to the encouragement of employees during those periods in the learning process when little or no improvement or accomplishment seems apparent to the learners themselves. For example, reference to Figure 12 on page 18 will show that the average looper in the hosiery industry

reaches what is called a *plateau* in learning from the seventh to the twelfth month on the job. During this plateau period little if any improvement in production is attained and as a result many employees, feeling that they have attained the limit of their production and being unsatisfied with their earnings, are likely to quit the job. A considerable investment is required by the company to bring an employee up to this level of performance. If an employee decides to quit at this point, the investment is largely wasted. One of the functions of the job trainer, then, is to keep the employees on the job until this plateau is passed. To do this requires a knowledge of the existence of the plateau and ability to explain it to the employees.

LEARNING TO LOOP

ACCEPT suggestions from instructor, or from experienced operators.
Assume that you are a learner until the end of the full period of six months (960 hrs.), especially during the last half of period try to correct faulty position, bad habits, unnecessary random movements, etc.
Below is a "looper curve," made from actual experiences of learners in Melrose.
Sometime after the 8th week (it may happen as late as 12th week) you will reach a "plateau," or leveling off period. This is to be expected. This period is very IMPORTANT, even though you do not gain dozens as fast as before.
There should be no absence from work during this time especially.
Lost time during any part of learning period delays progress.
Normally this plateau should not last longer than four weeks.
You will find that when you start to leave the plateau you will probably make greater progress than before you reached plateau
Try to make a dozen more today than yesterday
Hold yourself to a goal.
If work changes during learning period, you may lose dozens; do not become discouraged. Also some days you will make more than other days; the number of dozens will change slightly, but the general trend will be upward.
Do not let "making the code" be your goal; a good operator will greatly exceed the "code."
On 30 pt., learner will be slower during early weeks, plateau will come later, final dozens less. Confer with your foreman.

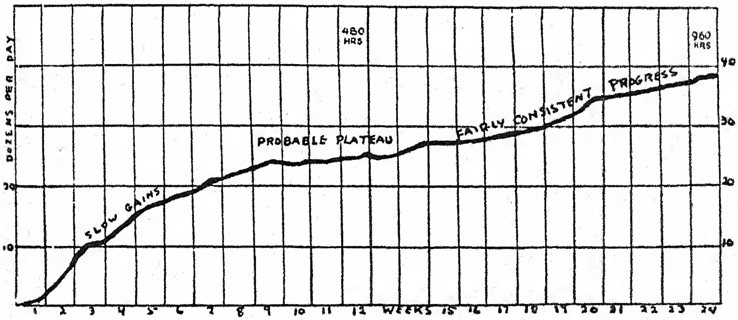


FIG. 99—Training material distributed to new loopers as a part of the training program of the Melrose Hosiery Company. (Courtesy of J. D. Boyd, Personnel Director, Melrose Hosiery, High Point, N. Carolina.)

One device used by the trainers in a hosiery plant[12] as an aid in teaching new loopers is illustrated in Figure 99 and Figure 100. The facts given on these cards, which are distributed to the new employees, serve to emphasize a number of points that facilitate learning the operation of looping.

PRACTICAL SUGGESTIONS FROM EXPERIENCED OPERATORS

Establish friendly, congenial relations with instructor and fellow workers. Everyone wants to help you succeed. Do not try to learn too fast. Be patient. Learn a little at a time

WHAT TO DO

1. POSITION. Try to relax. Some loopers urge that you sit erect, feet on floor with ankles crossed. Adjust chair to machine so that eyes meets points evenly.
2. HOLDING THE WORK. Hold **loosely** with thumb and index fingers of both hands. Some say use as many fingers as possible so as to hold lots of stitches in line that they may be pushed on with opposite fingers. The thumb is most important.
3. LEARN THE STITCH FIRST. Learn to put on stitches while machine is operating very slowly.
4. TURNING CORNER. Be sure to turn the first corner right and the last will be okay. Put on about third of sock, then turn corner by slipping stitches off with thumb and finger until directly in corner, then put top stitch over bottom stitch. When corner has been turned, continue to put on first side, then finish back side until complete sock is looped.
5. KEEP YOUR MIND on your work.
6. SEE HOW FAST you can pick up a sock after one sock is finished.
7. AFTER LEARNING, time yourself occasionally by dozens.
8. FORM THE HABIT, from the beginning, of inspecting your own work; at least two sock out of each dozen looped.

SOME DON'TS

Don't force work on, let it slip on "naturally"

Don't pull up or down, just be sure work is on the looping machine points "good."

Don't hold work tight; this tires hands.

Don't worry about how many dozens you should make, just keep at it.

Some say don't change hands when you turn corners.

Don't put the right hand underneath the looper; this is a waste of time. (Applies to right-hand loopers.)

Don't try to hurry; make speed by working steadily. Try to gain some each day.

Gossip, bad feelings, emotional upsets interfere with production. Insofar as possible everyone at Melrose wants good, congenial work conditions in the shop.

Fig. 100—Further training material distributed to new loopers as a part of the training program of the Melrose Hosiery Company. (Courtesy of J. D. Boyd, Personnel Director, Melrose Hosiery, High Point, N. Carolina.)

Job analysis for training

On page 28 several types of job analysis were defined. One of these is job analysis for training purposes. In this type of job analysis, a specific statement of steps to follow in performing the job is prepared for use of the trainer (either job trainer or supervisor). An example of this type of job

---

[12] Reproduced with permission of J. D. Boyd, Personnel Director of the Melrose Hosiery Mills.

INDUSTRY ________________ DEPARTMENT ________________

PAYROLL JOBS WITHIN DEPARTMENT

JOB TITLE	CODE	JOB TITLE	CODE
Drill Press Operator			

PURDUE INDUSTRIAL TRAINING MATERIALS
DIVISION OF EDUCATION AND APPLIED PSYCHOLOGY—PURDUE UNIVERSITY

TRADE & IND. FORM 1–12-42–1M

FIG. 101—Card for listing payroll jobs within the department. (Courtesy of C. H. Lawshe, Jr., Purdue University.)

PAYROLL JOB DRILL PRESS OPERATOR CODE 79.083

INDUSTRY ________________ DEPARTMENT ________________

UNIT OPERATIONS PERFORMED

1	Start and stop machine	11	Ream with press
2	Clean, oil, and care for machine	12	Tap with and without attachment
3	Adjust speeds and feeds	13	Tap with fixture
4	Set up work in vise	14	Counterbore and countersink
5.	Grind a drill	15	Drill at angle
6	Set up a drill (chuck, socket)	16	Multiple drill with attachment
7	Drill rough clearance holes through	17	Drill with offset head
8	Drill finish holes to size - through and to stop	18	Cut keyway
9	Drill with jig	19	Spot facing
10	Drill casting with core drill	20	Lapping

PURDUE INDUSTRIAL TRAINING MATERIALS
DIVISION OF EDUCATION AND APPLIED PSYCHOLOGY—PURDUE UNIVERSITY

TRADE & IND. FORM 2–12-42–2M

FIG. 102—Card for listing specific activities of a man on the job. (Courtesy of C. H. Lawshe, Jr., Purdue University.)

One device used by the trainers in a hosiery plant[12] as an aid in teaching new loopers is illustrated in Figure 99 and Figure 100. The facts given on these cards, which are distributed to the new employees, serve to emphasize a number of points that facilitate learning the operation of looping.

PRACTICAL SUGGESTIONS FROM EXPERIENCED OPERATORS

Establish friendly, congenial relations with instructor and fellow workers. Everyone wants to help you succeed. Do not try to learn too fast. Be patient. Learn a little at a time

WHAT TO DO

1. POSITION. Try to relax. Some loopers urge that you sit erect, feet on floor with ankles crossed. Adjust chair to machine so that eyes meets points evenly.
2. HOLDING THE WORK. Hold loosely with thumb and index fingers of both hands. Some say use as many fingers as possible so as to hold lots of stitches in line that they may be pushed on with opposite fingers. The thumb is most important.
3. LEARN THE STITCH FIRST. Learn to put on stitches while machine is operating very slowly.
4. TURNING CORNER. Be sure to turn the first corner right and the last will be okay. Put on about third of sock, then turn corner by slipping stitches off with thumb and finger until directly in corner, then put top stitch over bottom stitch. When corner has been turned, continue to put on first side, then finish back side until complete sock is looped.
5. KEEP YOUR MIND on your work.
6. SEE HOW FAST you can pick up a sock after one sock is finished.
7. AFTER LEARNING, time yourself occasionally by dozens.
8. FORM THE HABIT, from the beginning, of inspecting your own work; at least two sock out of each dozen looped.

SOME DON'TS

Don't force work on, let it slip on "naturally"

Don't pull up or down, just be sure work is on the looping machine points "good."

Don't hold work tight; this tires hands.

Don't worry about how many dozens you should make, just keep at it.

Some say don't change hands when you turn corners.

Don't put the right hand underneath the looper; this is a waste of time. (Applies to right-hand loopers.)

Don't try to hurry; make speed by working steadily. Try to gain some each day.

Gossip, bad feelings, emotional upsets interfere with production. Insofar as possible everyone at Melrose wants good, congenial work conditions in the shop.

FIG. 100—Further training material distributed to new loopers as a part of the training program of the Melrose Hosiery Company. (Courtesy of J. D. Boyd, Personnel Director, Melrose Hosiery, High Point, N. Carolina.)

Job analysis for training

On page 28 several types of job analysis were defined. One of these is job analysis for training purposes. In this type of job analysis, a specific statement of steps to follow in performing the job is prepared for use of the trainer (either job trainer or supervisor). An example of this type of job

[12] Reproduced with permission of J. D. Boyd, Personnel Director of the Melrose Hosiery Mills.

INDUSTRY ________ DEPARTMENT ________

PAYROLL JOBS WITHIN DEPARTMENT

JOB TITLE	CODE	JOB TITLE	CODE
Drill Press Operator			

PURDUE INDUSTRIAL TRAINING MATERIALS
DIVISION OF EDUCATION AND APPLIED PSYCHOLOGY—PURDUE UNIVERSITY
TRADE & IND. FORM 1–12-42–1M

FIG. 101—Card for listing payroll jobs within the department. (Courtesy of C. H. Lawshe, Jr., Purdue University.)

PAYROLL JOB DRILL PRESS OPERATOR CODE 79.083
INDUSTRY ________ DEPARTMENT ________

UNIT OPERATIONS PERFORMED

1	Start and stop machine	11	Ream with press
2	Clean, oil, and care for machine	12	Tap with and without attachment
3	Adjust speeds and feeds	13	Tap with fixture
4	Set up work in vise	14	Counterbore and countersink
5.	Grind a drill	15	Drill at angle
6	Set up a drill (chuck, socket)	16	Multiple drill with attachment
7	Drill rough clearance holes through	17	Drill with offset head
8	Drill finish holes to size - through and to stop	18	Cut keyway
9	Drill with jig	19	Spot facing
10	Drill casting with core drill	20	Lapping

PURDUE INDUSTRIAL TRAINING MATERIALS
DIVISION OF EDUCATION AND APPLIED PSYCHOLOGY—PURDUE UNIVERSITY
TRADE & IND. FORM 2–12-42–2M

FIG. 102—Card for listing specific activities of a man on the job. (Courtesy of C. H. Lawshe, Jr., Purdue University.)

analysis for a specific job is given in Table 21. This example is a specific break-down of "What to do" and "How to do it" for an operation of the United States Rubber Company.

A further illustration of materials that may be used effectively in making a job analysis for training purposes is given

UNIT OPERATION Rough drill clearance holes through CARD NO. 7
PAYROLL JOB Drill press operator CODE 6-78.083
INDUSTRY ________ DEPARTMENT ________

OPERATING STEPS	ESSENTIAL RELATED INFORMATION
1. Secure material	A. Familarize trainee with shop layout
2. Secure tools, lay out work	B. Explain tool and supply system
3. Center work according to instructions	C. Explanation of shop routine
4. Fit drill in sleeve	D. Safety precautions
5. Place drill and sleeve in spindle	1. General rules
6. Clamp up job in vise or strap clamps	2. Machine and accessories when rough drilling
7. Adjust table	3. Personal safety
8. Start machine	E. Lubrication of machine
9. Start machine in operation	F. Trade terms
10. Lower drill with hand lever to work	G. Blue print reading
11. Center up work, clamp to table, set stop	H. Speeds and feeds
12. Engage automatic feed, apply coolant	I. Selection of drill
13. Drill hole through	
14. Back up drill	
15. Stop machine	
16. Remove work	
17. Remove drill and sleeve from spindle	
18. Remove drill from sleeve	
19. Clean up press	
20. Remove burrs	

PURDUE INDUSTRIAL TRAINING MATERIALS
DIVISION OF EDUCATION AND APPLIED PSYCHOLOGY — PURDUE UNIVERSITY
TRADE & IND. FORM 3-12-42-5M

FIG. 103—Card for listing the detailed steps and related information necessary for performing each part of the job. (Courtesy of C. H. Lawshe, Jr., Purdue University.)

in Figures 101 to 103. Three series of these cards are prepared. On an A Card, illustrated in Figure 101, are listed all payroll jobs in the department. For each of these jobs a B Card, illustrated in Figure 102, is prepared. On this card is listed every operation that is performed by a man on this job. For each operation a C Card, illustrated in Figure 103, is prepared. When the analysis is complete for any given job, there will be as many C Cards for that job as there are operations to be performed. The C Cards thus become the training manual for the job. A careful use of this kind of analysis in training new employees avoids many omissions in the training that are almost certain to be made if some systematic procedure is not followed. An application of this

TABLE 21

HOW TO PREPARE THE STRIP OR FOXING CALENDER FOR OPERATION

Step	*What to Do*	*How to Do It*
1.	Advise mill man of first compound required	*a.* Check ticket. *b.* Indicate type and approximate quantity of run.
2.	Advise take-off crew of type of material being processed.	*a.* Check list. *b.* Indicate type and approximate quantity of run. *c.* Indicate whether boards or rolls will be used.
3.	See that correct engraved roll is in calender.	*a.* Check list. *b.* The #4 roll is interchangeable to obtain correct engraved rubber surface. *c.* See breakdown titled "How to Change a Calender Roll."
4.	Apply heat on all four rolls.	*a.* Be sure couplings are fastened. *b.* Open individual valves for each calender roll. *c.* More steam on rolls #3 and #4. *d.* Less steam on rolls #1 and #2.
5.	See that all moving parts are free.	*a.* Be sure static discharges are out of way. *b.* Be sure no foreign object is near the bite of the calender rolls on the drive gears.
6.	Start machine and test safety devices.	*a.* Start motor (usually a separate button). *b.* Push starting button on central panel. *c.* Decrease speed to slowest speed—as soon as calender starts. *d.* Pull safety stop cable from both sides of machine. *e.* Test stop button on control panel.
7.	Apply heat on next roll required.	*a.* Check ticket. *b.* Locate next roll on storage rack. *c.* Apply heat at storage rack.
8.	Place cutting bar on calender (if running strips or foxing).	*a.* Check ticket. *b.* See that correct cutting bar is in place on the calender above roll #6. *c.* See that cutters are in "no cut" position, "dog" released and lever up.
9.	Adjust boards at back of feed rolls #1 and #2.	*a.* Set slightly narrower than desired width of finished calendered stock.

TABLE 21—*Continued*

Step	*What to Do*	*How to Do It*
10.	See that rolls are clean.	*a.* Clean surface with gasoline and cloth. *b.* Clean excess grease and oil from edges of rolls. *c.* Be sure all rubber scrap removed from crevices at edges of rolls.
11.	Check to see that calender crew is ready for run.	*a.* Warming mill man has proper stock at correct temperature. *b.* Take-off crew has correct reels or boards.

method of job analysis for a specific operation is given in Table 21.[13]

Another type of job analysis for training purposes is illustrated by the work of Lindahl,[14] who set up a simple graphic recording apparatus to show the stroke pattern

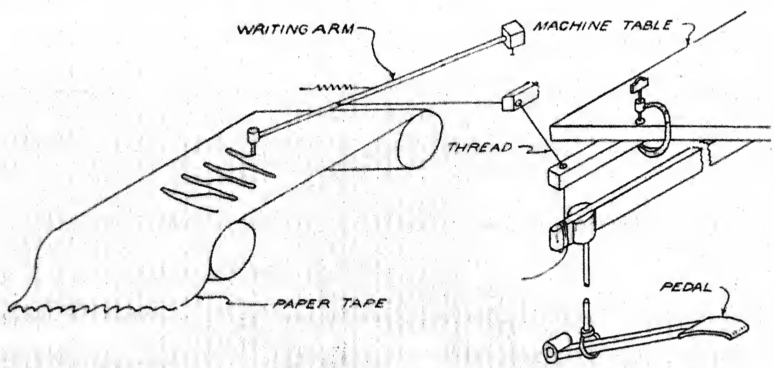


FIG. 104—Schematic drawing of apparatus used in obtaining graphic records of job performance in cutting discs with a foot-operated abrasive wheel.

followed in cutting tungsten discs. The machine studied made use of a rotating abrasive wheel that is lowered by means of a foot pedal to cut off a disc. After the disc has been cut, the foot pressure is released, the abrasive pressure is relieved, the abrasive wheel rises, the tungsten rod advances, and the

[13] By permission of the Personnel and Training Section of the United States Rubber Company, New York, New York.

[14] Lindahl, *op. cit.*

machine is ready for a repetition of the cutting cycle. The recording apparatus used by Lindahl is illustrated in Figure 104. Identification of the various parts of the cutting cycle is shown in Figure 105.

The cutting patterns found at various stages of training in the performance of a typical trainee are illustrated in Figure 106. By using this simple equipment to show trainees the specific errors they were making in their cutting cycles, Lindahl was able to improve the quantity and quality of work at various stages of training, and also to decrease the use of abrasive wheels.

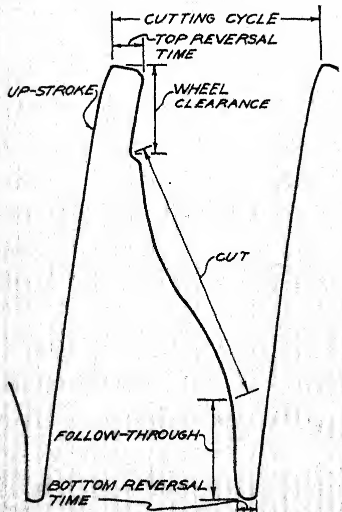


FIG. 105—Identification of various parts of the cutting cycle as shown by the graphic record attained with the apparatus shown in Fig. 104.

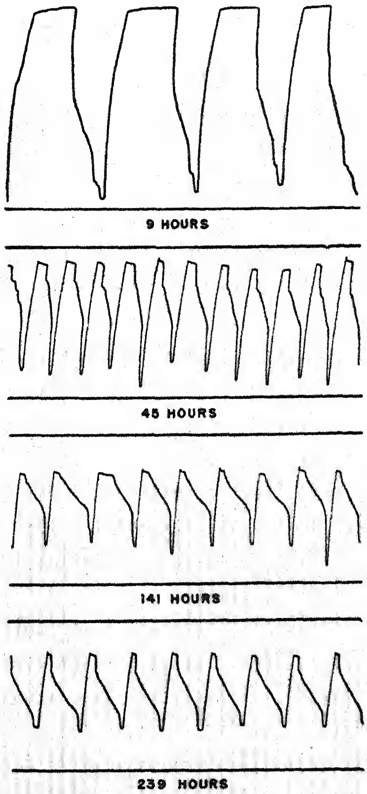


FIG. 106—Records of performance of a trainee at various stages of his experience on the job.

Other training techniques

A technique used by the armed forces with great success during World War II is the tachistoscopic airplane recognition

training. This approach, founded upon the basic work of Renshaw,[15] involves presentation of stimuli for a very brief period (often $\frac{1}{10}$ of a second or less). By use of this method, subjects have been taught to see and identify distinguishing features of objects or forms when the "flash" presentations are reduced to an exceedingly small fraction of a second. In fact, Renshaw found that when trained by this method ordinary people rapidly achieve a degree of skill heretofore considered genius. The application of this approach to aircraft recognition training has been described by Terrell.[16] Although this method of training has apparently not been used extensively in industry, it would seem to offer real possibilities for certain types of training, particularly on such jobs as inspection for appearance.

The use of films, slides, and other visual aids is another technique that has been used successfully in certain types of industrial training. VanderMeer[17] reports a study in which sound films were used in the training of engine lathe operators. He found a significant superiority in the operators trained by this method when compared with those trained by traditional methods.

The job trainer has many responsibilities in addition to instruction in the specific routine of job performance. Often he is the only man with whom the new employees are in daily contact. The impressions of the company which they form in the early stages of their employment, and which often remain with them as long as they remain employees, are often largely dependent upon their reactions to the job trainer. He must be able to explain company policy to the employees

---

[15] S. Renshaw, "The Visual Perception and Reproduction of Forms by Tachistoscopic Methods," *Journal of Psychology*, XX (1945), pp. 217–232.

[16] C. M. Terrell, "Recognition Training in the Army Air Forces," *Military Review*, Fort Leavenworth, XXIV (1945), pp. 40–43.

[17] A. W. VanderMeer, "The Economy of Time in Industrial Training: An Experimental Study of the Use of Sound Films in the Training of Engine Lathe Operators," *Journal of Educational Psychology*, XXXVI (1945), pp. 65–90.

and to give them convincing reasons for the existence of these policies. He must be able to explain the wage-payment plan in operation and to show why this plan has been adopted by the company rather than some other plan which the employees might think is fairer. If he finds that certain phases of company policy are not in general understood or are not accepted by new employees, it is his job to relay this information to the department of industrial relations or to the other management men who are responsible for the adoption of the policies in question.

Employee training, as a branch of management, is becoming so important that many industries have set up as a part of their industrial relations department a specific training division.

Supervisory training

Perhaps no job in industry has undergone such great changes during the past twenty years as that of the supervisor or foreman. Two decades ago, when a supervisory position was to be filled, the best producer in the department was usually given the job without regard to his ability (or lack of it) in the industrial-relations phases of the supervisory position. As a result of this policy the foreman was often a "bull of the woods." He hired and fired, set the rates for his men, dealt with problems of discipline as he saw fit, was responsible for his own quality control, determined methods of work, did his own training, and took care of his own maintenance. The emergence and development of scientific management gradually took many of these responsibilities away from the foreman. Problems of hiring have largely been taken over by the personnel department. Responsibility for layoff has been assumed by the department of industrial relations. Wage rates have been established on a systematic basis by means of job evaluation (see pages 374 to 393). Problems of discipline have largely been delegated to the

department of industrial relations. Quality control has often become so dependent upon the use of technical equipment that the creation of an inspection department has been necessary. Methods of determining the most efficient layout of work have been taken over by motion study specialists from the department of industrial engineering. Because of the increasing complexity of machinery, a specialized department dealing with the maintenance of equipment has been established. And now the training of new employees is becoming specialized with job trainers assigned to do work that was formerly conducted by supervisors alone.

Many foremen feel that these changes have gradually removed from the supervisor so many of his former duties and responsibilities that the very need for the foreman has almost disappeared. But with the removal of many of his former responsibilities and duties has come the assignment of many new duties with which the foreman was not concerned twenty years ago. The foreman of today must know the law. Daily, almost hourly, problems arise and decisions must be made in which the voice of the foreman is the voice of management. Recent social legislation has made it imperative that decisions made by management with respect to such matters as layoff, transfer, seniority, hours of work, wages, and a host of other matters must conform to certain legal requirements. Decisions on such matters must be made by the foreman. Often these decisions must be made at once, with no time for conferences or for consultation of statements of the law. Decisions may involve such problems as hours of work, settlement of a grievance, or interpretation of a union contract for determining the seniority of one man over another. Such decisions cannot be made by rule of thumb. They must be made, always, in the light of written laws, policies, and union contracts. The foreman must know these documents and their interpretation. He must be able quickly and accurately to size up a situation and outline a course of action that will not be in conflict with laws or agreements. A good producer

might have made a good supervisor twenty years ago, but today the foreman must in addition be something of a lawyer and, incidentally, a pretty good one.

The foreman today must be a statesman who has both wisdom and power to act. Disputes were settled easily twenty years ago. The rule was simple and easy to learn—in any dispute the foreman was right and the employee wrong. Today both may be right or both may be wrong, or each may be partly right and partly wrong. A decision reached by management today must be made after consideration of all angles and the viewpoints of all parties. It is a commonplace observation that when authoritarian judgments give way to discussion as a means of settling grievances, a degree of mutual understanding must be developed by both parties. The foreman of today must not boss his men so much as he must understand them. He must know why this one becomes disgruntled upon slight pretext, why that one is often late to work. His job is to help the men help themselves. This is particularly true in periods of emergency production when, as one foreman put it, "You don't dare bawl a man out. If you do, he might quit." The foreman of today must know how to handle men without making use of a "big stick." Though the foreman has seldom studied the subject of psychology, it is very important for him to know, if he is to do his job well, that the desires and wishes of his men must be both understood and considered. Without such understanding he cannot hope to be a satisfactory supervisor.

The foreman must look upon the experts in personnel, motion study, rate setting, and so on, not as rivals and substitutes but as specially trained men who have been assigned to help him with his job. He must understand the services performed by these experts and be able to explain these services to the men. The wage incentive plan may be installed by an industrial engineer, but it is the foreman, not the industrial engineer, who must be able to explain the plan to the employees. When an employee cannot figure out why

his pay check came to only $48.60 instead of $48.70, the foreman must be able to explain how the figure of $48.60 was reached. If he cannot do so, the worker may be disturbed out of all proportion to his imagined loss of ten cents.

Changes in the job of the supervisor such as those described above have made it necessary for management to devote considerable time and effort to the training of supervisors. Many industries have set up a complete program of supervisory training conferences, the purpose of which is to train the present supervisors in the newer aspects of their job that are developing almost daily and to prepare younger men who are to become supervisors. One plan in common use is to have every supervisor, or potential supervisor, attend a conference or class on company time for a period of from one and a half to two hours at approximately two-week intervals. The subject matter covered in such conferences varies considerably from plant to plant, and in each case is decided upon by management in the light of the needs of the company. A typical list of the subjects covered in one series of supervisory conferences is shown in Table 22.

While many of the topics listed in Table 22 deal with matters of production, it will be noted that several are devoted to such matters as company policy, industrial relations, union contracts, and the services of experts in such fields as job analysis and insurance.

It has already been emphasized that some means of measuring the effectiveness of training should be employed. A typical method of measuring the effectiveness of training in company policy is illustrated by supervisory case problems, with related true-false questions, such as the one reproduced in Table 23 on pages 280–281.

The Joe Blake case (the name is fictitious) is an actual case that arose in one plant during one of the conferences on company policy. The answers to the twenty true-false questions obtained from a group of supervisors not only furnish management with a means of knowing how well prepared

TABLE 22

TOPICS COVERED IN A SERIES OF SUPERVISORY TRAINING CONFERENCES[18]

1. Accident Prevention
2. Waste Prevention
3. Training for the New Employee
4. Re-training of the Present Employees
5. Cost of Maintenance
6. Cost of Materials Used in Ingot Production
7. Incentives (Financial and Non-financial)
8. Job Training
9. Stability of Employment
10. Co-operation
11. The Cost of a Ton of Steel
12. Direct Application of a Cost System
13. Inspection
14. Recognition of Ideas (old and new)
15. Transportation (inter and intra)
16. Methods of Handling Complaints
17. Plant Protection
18. Why Condition—Semi-finished Material
19. Discipline
20. Proper Methods of Dealing with Men
21. Product and Machine Innovations
22. Publication of Management's Policies (uniformity)
23. The Application of Conference Conclusions
24. Job Analysis versus Man Analysis
25. Service—What is it?
26 Quality versus Quantity
27. Seniority
28. "What's on the Worker's Mind?"
29. Collective Bargaining
30. Unemployment Insurance

supervisors are to handle this type of problem but also indicate what areas of company policy should be covered in the supervisory training conferences in order to achieve a unified understanding of company policy on the part of supervision. A series of twenty cases, with twenty true-false questions on each case—from which Case 1 was taken—furnishes management with a 400-item true-false test with which to measure how much the supervisors know about the administration of

[18] R. J. Greenly, "Conference Leadership: Report on First Conference Leaders' Conference, Carnegie-Illinois Steel Corporation, Gary Works." (Indiana State Board for Vocational Education in Co-operation with the Training Division of the Industrial Relations Department, Carnegie-Illinois Steel Corporation, Gary, Indiana, 1937.)

TABLE 23

A Test Used in Measuring What Supervisors Know About Company Policy*

CASE NUMBER 1

Plant ....

Dept. ....

CASE NUMBER 1
SUPERVISORY PROBLEMS

Joe Blake is a tractor operator in No. 8 shipping department. He has worked for the company six months and had a clean record until a week ago. Before coming to this job he worked for a neighboring company and was highly recommended by them. A week ago his tractor became overheated because of a dry radiator and had to be sent to the shop for repairs. Both shipping and repair departments were behind schedule and the resulting delay due to the absence of this tractor caused considerable confusion and loss.

One of the tractor rules states that each operator shall inspect his machine thoroughly before work is begun, and also before turning the machine over to the next operator. Joe Blake, like all other operators in the department, does not make the first inspection but, according to agreement among the tractor operators, makes a careful inspection at the end of the turn so the following man finds the machine ready to go.

After investigating, the foreman reported the incident as an act of carelessness and a clear violation of the plant rules, and consequently penalized Joe by giving him a one-week lay-off. Joe reminded the foreman that it was a well-known and tolerated practice to omit the first inspection and brought in the previous operator who insisted that he had filled the radiator just before the breakdown. Being a new man, Joe did not feel it wise to argue further and accepted the penalty with good grace.

During the following week the rumor arose and was substantiated that the tractor breakdown was caused by a group of six packagers who, as a joke, drained the radiator of Joe's tractor while he was away in the lavatory. These details and the names of the men involved were known to nearly everyone in the department.

INDUSTRIAL RELATIONS—CARNEGIE-ILLINOIS STEEL CORPORATION

The following statements apply to this situation. Study each one carefully. If you agree, put a circle around the "T" following the statement. If you do NOT agree, encircle the "F." If you are in doubt, encircle the double "??." Be sure to show your response clearly.

1. The six packagers should be given penalties equal to Joe Blake's.	T	??	F
2. The foreman should apologize to Joe and void the rule violation report.	T	??	F
3. The foreman should let the report stand as it is, since responsibility for the continuous operation of the tractor is Joe's, regardless of the action of others.	T	??	F
4. Joe should be given enough overtime work to make up for his week's loss.	T	??	F

5. The foreman should be fired for allowing the inspection rule to be violated openly. T ?? F
6. The foreman should not change his action unless Joe takes the matter up with his grievance committee and threatens to demand redress. T ?? F
7. The departmental supervisor should take advantage of the situation by holding a meeting with the men, pointing out the waste and loss caused by a "practical joke" and a violation of plant rules. T ?? F
8. Joe should be transferred to another department and other steps should be taken to "hush up" the foreman's action. T ?? F
9. The company should refuse to recognize the act of the packagers since delayed recognition of that aspect of the case might be held against the company at a later date as an example of management's arbitrary policies carried out without proper investigation. T ?? F
10. A general survey should be made to see that plant rules are being obeyed to the letter throughout the department. T ?? F
11. All the tractor operators in the department should be penalized for wilful neglect and carelessness because of their practice in regard to the inspection rule. T ?? F
12. The foreman should protect his men by explaining to his supervisor that he had known of and tolerated the limited inspection practice. T ?? F
13. The foreman was not responsible for the limited inspection practice even though he knew of the rule violation. T ?? F
14. The foreman should conceal the matter from his supervisor, if possible, since admission of his tolerance of a rule violation would serve only to increase friction and undermine his authority. T ?? F
15. The company should not permit the foreman and superintendent to admit that an error of hasty misjudgment was made, since that would lower employee morale. T ?? F
16. The tractor operators as a group had no responsibility in this case because the inspection practice was tacitly approved by the foreman who is management's representative. T ?? F
17. The responsibility for the occurrence rests solely with the six packagers. T ?? F
18. This is one of those situations in which the responsibility cannot be fixed. T ?? F
19. Blame for the occurrence rests solely with the foreman. T ?? F
20. A foreman should be certain of his information and of all the facts before he takes action against an employee. T ?? F

---

* R. J. Greenly, "The Case Method in Supervisory Training," American Management Association Personnel Series No. 56, (1942), pp. 30–35.

company policy. True-false or objective questionnaires, formerly considered purely academic in application, have been found to be serviceable as measuring instruments in practical industrial training situations. Similar use may be made of multiple-choice questions dealing with company policy.

Apprentice training

The training of operators—either old or new employees—can ordinarily be carried on as the need for operators arises. An immense number—perhaps a large majority—of industrial employees are engaged in semiskilled jobs that require an intensive but not an extensive period of training. Trainees for such jobs can often be prepared in a relatively short period of time—often only two or three weeks. The training of supervisors, on the other hand, requires considerably longer. An adequate understanding of the increasingly complex duties of the modern supervisor ordinarily requires a several months instructional period. Industries must therefore use considerable foresight in anticipating supervisory needs so that training programs may be undertaken in advance of the actual need for supervisors.

Another area of industrial training in which still greater foresight must ordinarily be exercised is the training of skilled tradesmen. This training is usually accomplished by means of an apprentice program. Except for college-trained technical men there is perhaps no other single group of industrial employees who require so extensive a period of training for expert performance as do skilled tradesmen. The federal committee on apprenticeship training specifies that the training period for apprentices shall be not less than 4,000 hours, and that the shop instruction shall be accompanied by not less than 144 hours of instruction under public authorities.[19] These stipulations reflect the well-substantiated judg-

[19] The National Apprenticeship Program (War Manpower Commission Bureau of Training, Apprentice-Training Service, 1943).

ment of those experienced in this field that adequate training of skilled tradesmen requires an intensive and extensive instructional period. These facts make it more necessary for industry to look into the future in the training of skilled tradesmen than in the training of many other types of industrial employees. The various details in the administration of an apprentice program, together with the contractual agreement between the apprentice and the employer, are not of primary concern to the technical psychologist, and therefore will not be discussed here. Psychology, however, may be definitely applied in the use of aptitude and qualification tests for the selection of boys to receive apprentice training. While mass production in industry has reduced the proportion of employees who must be skilled tradesmen, the design, construction, and maintenance of machines has increased so much in complexity that a very high level of capacity is required of the boy who is to become a skilled tradesman. Figures 5 to 8 in Chapter 1 show that the range of ability among applicants for apprenticeships is large indeed. Any program of apprentice training that does not provide an adequate means of selecting trainees according to their capacity to benefit from the training is likely to result not only in a failure to develop expert tradesmen but also in the waste of a considerable company investment in the training given. Those responsible for the administration of an apprentice program should therefore work in close collaboration with the testing division of the personnel office so that trainees will be selected who are able to take full advantage of the opportunity offered.

Other fields or types of training

A complete industrial training program provides for several types of training in addition to those mentioned above. Among these is the training of college graduates. It has long been recognized that the average college man needs help in bridging the gap between the campus and the plant. Every

industry has specialized equipment, processes, and methods of production that are quite unknown to the college student at the time he goes to work. A training program that permits him to spend a certain amount of time in each of several departments so as to obtain a background of practical experience which he can relate to his technical training is an indispensable part of an adequate training program for newly inducted college men. This problem is met to some extent by the so-called co-operative systems of higher education in which the student alternates between college and industrial employment, each activity occupying approximately a six-months period. This procedure does not entirely solve the problem, however, partly because the co-operative plan is not in extensive use and partly because even where practiced it acquaints the student only with those industries in which he has had his work experience. Plants differ markedly not only from one industry to another but even from one plant to another in the same industry. Until a student has had an opportunity to learn at first hand the production problems of the plant in which he is later to work, he cannot be expected to be a very valuable member of the organization. Thus, special training for college men is nearly always necessary as one phase of a training program.

Related trade training in various forms offered by public-school systems is another type of training commonly used by industry. With the passage of the Smith-Hughes Act in 1917 and the George-Deen Act in 1936, certain funds were made available to school systems through the state offices of education to encourage the development of various forms of vocational education. The particular type of education or training that may be offered by a public-school system can usually be determined largely by the needs of local industries, if the industrial men responsible for the training are willing to work in co-operation with the school officials in organizing the courses of study. Under this plan, school systems may give in evening and extension classes technical training in

areas related to the jobs of men employed by local industrial plants. This system not only removes from the industry a considerable responsibility in the training of employees in newer developments of their work, but also furnishes a desirable relationship between the schools and the industries.

Correspondence courses furnish another type of training that is often helpful in keeping employees abreast of their jobs and technically prepared to assume more responsible positions. However, certain objections to correspondence courses as a means of training have been raised. These courses, though usually of considerable value in fields that have been well standardized, often do not take into consideration the specific needs of any local industrial situation. It is also true, unfortunately, that often the employees who elect to take the correspondence courses are not the ones who would profit the most from them. As Bird and Paterson[20] have pointed out, correspondence schools could be of much greater service if, instead of selling a course to anyone, they would select students in terms of aptitude and previous training. As it is, only 6 per cent of the group of 305 persons investigated by Bird and Paterson completed the course which they had started.

Some Characteristics of Learning

The problems of industrial training may properly be considered as a branch of educational psychology or the psychology of learning. A great many investigations dealing with the problems of learning have been conducted; from these investigations a set of general principles, part of which are applicable to the industrial situation, has emerged.

The learning curve

Almost every psychological investigation of a learning activity results, among other things, in the establishment of a

[20] C. Bird and D. G. Paterson, "Commercial Correspondence Courses and Occupational Adjustments of Men," *Bulletin of the Employment Stabilization Research Institute, University of Minnesota*, II (1934), 27 pp.

learning curve. Such a learning curve for the operation of looping in the hosiery industry is reproduced in Figure 12 on page 18, and a similar curve is reproduced on a card distributed to new employees, as illustrated in Fig. 91. Learning curves for the operations of trimming, covering, and hemming in the textile industry are reproduced in Figure 90 on page 251.

The learning curve for a given operation shows at a glance the level of skill that has been reached by the average operator after a given period of training or experience on the job. Anyone who is responsible for the training of new employees on a specific job will find that a learning curve obtained from the records of previous employees is of considerable value. From such a curve he will be able to see immediately whether the progress of an operator is up to standard and, if not, how far the performance is below standard at any time in the instructional period. He will also be able to detect immediately whether an employee is maintaining his former level in comparison with the standard curve or is gaining or losing in comparison with other employees. In short, he will have a basis of comparison by means of which he can evaluate the effectiveness of his own instruction and the effectiveness of the learning that is being done by the new employees. If an employee falls below the curve by a significant amount, and, in spite of every effort of the employee or the instructor to remedy the situation, continues to lose ground in comparison with the curve for the average, it is possible that the employee should be transferred to some other job before too much time and effort have been wasted in training him. It is even desirable, in some instances, to supply each new employee with a copy of the learning curve for the job and allow him to plot his own curve on the same chart, thus giving him a graphic record of his progress from day to day or week to week.

It should not be inferred from the foregoing discussion that there is any generalized form of learning curve that can be reduced to a mathematical equation and that applies to all learning situations. Many factors determine the form of a

learning curve, and these factors vary greatly from one situation to another. Since one of these factors is the method of training, it is not uncommon to find different learning curves for the same industrial operation when different methods are involved. Indeed, a comparison of the resultant learning curves for different methods of training is an excellent way to compare the effectiveness of the methods. But with the various factors, such as method of training, held as constant as possible, a reasonably standardized curve results for a given operation for the average person. It is this learning curve that the job trainer should have available if he is to evaluate adequately the effectiveness of his instruction and the progress of the trainees who are under his supervision.

The plateau

In learning any complex task it often happens that, after a certain level of efficiency has been attained, a period of time arrives in which little or no improvement takes place. This period is followed by a later increase in skill. The period during which no apparent improvement occurs is known as a *plateau*. Plateaus are not always present in the learning curves for industrial operations, and when they are present it is sometimes possible to eliminate them by changed or improved methods of instruction. No matter what method of instruction may be in use, however, arrival at a plateau in the learning process often is characteristic of the average learner's performance.

A typical example of a plateau in a learning curve is shown in Figure 12 on page 18. We do not need to concern ourselves here with the cause of the plateau, although it might be mentioned that a considerable amount of psychological investigation has been devoted to this subject and a number of experiments have been made which have identified certain factors that may result in a plateau.[21] If a plateau is characteristic of the average learning curve of an industrial operation

[21] W. H. Batson, "Acquisition of Skill," *Psychological Monographs*, XXI (1916), pp. 1–92.

it is highly important for the job trainer who is supervising this training to be aware of the situation. Without such knowledge he may feel that his training is inadequate, that the trainees have lost interest or that they have reached the maximum level of which they are capable and therefore need no further training, or that many of the employees should be shifted off the job because they are unable to reach the level of production achieved by experienced operators. If, on the other hand, he is aware of the existence of the plateau, he will continue his instruction, encourage the trainees to stay with the job until the end of the plateau is reached, and hold in abeyance his final judgment of the trainees' ability until they have had an opportunity to benefit from the final spurt in production that is certain to follow the plateau if the training is continued.

By analyzing the activities occurring during the plateau period it is often possible to modify the training program so that the plateau may be partially or entirely eliminated. Whether or not the job trainer is able to achieve this result, his knowledge of the existence of the plateau, together with his awareness of the fact that this situation is by no means uncommon in the learning of industrial tasks, will enable him to cope with the situation in a more intelligent manner and to help tide the employees over a period that, at best, is likely to involve a certain amount of discouragement.

Knowledge of results

One way in which training may be, and often has been, given is briefly to explain the work to the new employees and then to "turn them loose," so to speak, to perform the job without systematic or adequate check-up on the quality of their work. Another method is to follow the initial instruction period by frequent and systematic study of job performance and to provide a means for informing the trainees or new employees in exactly what respects they are doing the job correctly or incorrectly. Both experience and experiment

have shown that the former method of training individuals in any new task results in a considerably lower level of job performance than the latter method. This principle was first pointed out in connection with the learning of arithmetic among the school children. Panlasigui and Knight[22] found that if children are consistently informed of all errors made in arithmetic they show a significantly more rapid improvement in learning the arithmetic computations involved than if they are given exactly the same sets of drills and are not provided with a "knowledge of results." Book and Norvell[23] found the same principle operating in the learning of mental multiplication. In their experiments the persons in the experimental group watched the record of their progress and attempted at each practice to make a higher score than the score made previously. Persons in the other, or control, group were not informed of their progress. These results are summarized graphically in Figure 107. After the tenth practice period the conditions were reversed, so that the persons formerly receiving knowledge of results were no longer informed of progress. The persons formerly uninformed of accuracy were now given information as to their progress. The curves in Figure 107 show that immediately after the tenth practice the groups changed their relative standing. The group that was given knowledge of results forged rapidly ahead of the other group.

One need not look long in industry to find situations parallel to this experiment. The work of Lindahl[24] previously mentioned is an example of the application of the principle of knowledge of results to an industrial learning situation. In the investigation of inspectors summarized in Chapter 9 the accuracy of many inspectors was found to be significantly

[22] I. Panlasigui and F. B. Knight, "The Effect of Awareness of Success or Failure," *Twenty-Ninth Yearbook of the National Society for the Study of Education*, Part II (1930), pp. 611–619.

[23] W. F. Book and L. Norvell, "The Will to Learn," *Pedagogical Seminary*, XXIX (1922), pp. 305–362.

[24] Lindahl, *op. cit.*

below that of which they were capable and also below the level of accuracy required for satisfactory job performance. It was found that the unsatisfactory level of accuracy characteristic of many employees in this department was traceable to the training procedure that had been followed prior to this experimental work. The former training method consisted in placing each new operator between two older operators who

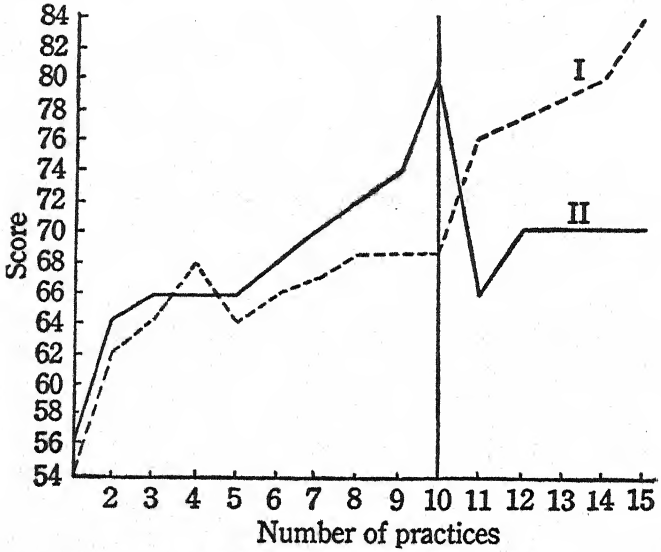


Fig. 107—How knowledge of results affects performance in learning a new task.

were asked to instruct the new employee in the details of the job. Since each of the older employees had a definite amount of work to perform during the day it was only natural that they should spend little if any time in training the new employee. Indeed, it was found in many instances that the older employees resented this additional training load and as a result did not give the new employee adequate training in the details of the inspection job. This method of training resulted in a lack of standardization of inspection practice among the employees—many of the inspectors were over-

looking certain defects that they were *able* to detect but which they had not been trained to observe. The new training program instituted in this department included the assignment of specific job trainers to the task of teaching new employees, so that the latter would be able to develop satisfactory standards of performance without guessing at which sheets were satisfactory and which were defective.

The general principle of job training is that new employees will improve very slowly—if, indeed, they improve at all—unless they are provided with systematic and accurate information on the quality of their work. Provision for furnishing new employees with definite knowledge of results should be an integral part of any training program.

Motivation

Ordinarily speaking, people will not learn very much about anything unless they are motivated to do so, that is, unless they are supplied with an adequate incentive. Incentives in industry are usually considered under two general heads: financial and nonfinancial. Under financial incentives are considered those that pertain to methods of wage payment, such as an hourly rate *versus* some form of piece work. One of the major problems in industrial management is that of deciding upon a wage payment plan that will result in a satisfactory level of job performance and at the same time will be considered by the employees to be a fair plan. If the plan in use does not place a proper premium upon high production, it often happens that the employees never bother to learn the skill required for a very high level of job performance.

A typical example of this situation is furnished in a study by Kitson.[25] Kitson examined the production records of 40 hand compositors whose experience ranged from one to

[25] H. D. Kitson, "Extra Incentive Wage Plans from a Psychological Viewpoint," *Bulletin American Management Association, Production Executive Series*, No. IX (1925), p. 7.

twenty-seven years, the average being eight years. The average output of these employees was 55 units for a week of work on a scale of efficiency established by time studies. An incentive system of wage payment was installed which permitted the employees to earn a bonus for production over the previous average amount. Within five months these employees were averaging 97 units on the scale and at the end of a year the average output had reached 103 units. This experiment made it clear that the previous method of wage payment did not provide sufficient motivation for these employees to learn the skill necessary for high production, even though they had been on the job a number of years.

Incentives are not limited to financial incentives. A number of investigators of this subject, for example Lee[26] and Miles,[27] have stressed the fact that such factors as recognition of good work, a fair system of promotion, and job security all lead to an attitude of mind on the part of the employee that helps develop a desire to learn the job well and to turn in a satisfactory performance on the job.

Another aspect of motivation upon which the evidence is reasonably clear and convincing is the relative desirability of rewards and punishments. Translated into industrial terms, this question raises the issue as to whether a new employee will learn faster and better if he is complimented for all good job performance or if he is criticized for all poor job performance. Since any new employee will do some things correctly and other things incorrectly, it is quite possible to use either or both of these methods while he is learning the job. Judging from a number of related investigations in the field of experimental psychology, it seems clear that praising satisfactory performance is the more desirable method to follow.[28] Mod-

---

[26] C. A. Lee, "Some Notes on Incentives in Industry," *The Human Factor*, VI (1932), pp. 180–186.

[27] G. H. Miles, "Effectiveness of Labour Incentives," *The Human Factor*, VI (1932), pp. 53–58.

[28] E. L. Thorndike, *Effects of Punishment and Reward* (University of Chicago Press, 1932).

ern industrial practice has largely adopted this point of view, primarily because it results in more satisfactory industrial relations.[29] It is interesting and fortunate that experimental studies in psychology substantiate this judgment for the quite different reason that praise is superior to criticism as a means of inducing the employee to learn.

Attention to specific features of the job

All but the simplest of industrial jobs require several different types of skill or ability. The stenographer must be able not only to take shorthand but also to use a typewriter. The inspector is often required to spot certain defects of appearance visually, other defects by weight or feel, and perhaps still other defects by means of readings taken from special instruments. It often happens that the several aspects of a job are sufficiently independent of one another to necessitate the giving of special training for each. Reference to Table 25 on page 305 shows that the intercorrelations between accuracy of inspectors in detecting different kinds of defects are in general quite low. This means that the inspectors who are best in detecting certain defects are not necessarily better than average in detecting other defects. To learn an inspection job efficiently new inspectors cannot simply be told to "throw out the defective material." They must be given specific instruction in the nature of the defects and taught what to look for in each specific case. Training must be specific and to the point instead of haphazard and general.

Transfer of training

Does training on one industrial job transfer to another job? Will an employee who has been thoroughly trained in one operation be easier to train in a new operation than an employee who has not had this previous training? These

[29] J. J. Jackson, "Reprimanding Employees," *Personnel Journal*, XIX (1941), pp. 73–80.

questions have stimulated a great deal of investigation by psychologists. At one time it was felt (with little or no experimental evidence to support the feeling) that training would develop certain general "powers" or "faculties" so that a person trained in one kind of activity would be more able to do many other kinds of activities in which these faculties might be used. It was felt that the memorizing of Latin verbs and conjugations would improve one's memory so that he would be better able to memorize insurance premium tables or grocery price lists. In industrial terms, training in the operation of one machine, it was felt, would improve an employee in ability to operate some other machine with which he had had no direct experience.

This notion—the wholesale *transfer* of *training*—has been subjected to a great deal of experimental work, beginning with that of Thorndike and Woodworth.[30] The conclusion reached by these early investigators—and substantiated by several more recent studies—is that the amount of transfer from one skill or mental function to another is far less than was formerly believed. Indeed, Thorndike[31] has concluded that transfer in a general way does not occur at all and that what is often regarded as transfer is simply due to *identical elements* in the two jobs under consideration. These identical elements may be activities (usually manipulative activities in industry) or methods of work. An example of the former would be any routine assembly operation that is identical with the operation in some other assembly job. An example of the latter would be the methods of laying out work that would be similar from one job to another in the activities of a motion study or layout man.

The conclusion to be kept in mind is that supervisors

---

[30] E. L. Thorndike and R. S. Woodworth, "The Influence of Improvement in One Mental Function upon Efficiency of Other Functions," *Psychological Review*, VIII (1901), pp. 247–261, 384–385, 553–564.

[31] E. L. Thorndike, "Mental Discipline in High School Studies," *Journal of Educational Psychology*, XV (1924), pp. 1–22, 83–98.

should not expect more transfer to occur than can be expected in terms of definite similarities in the jobs. Because an employee has been able to learn and successfully perform one job is no guarantee that he will do as well on some unrelated job. On a second job he may succeed or he may fail, but his performance is neither limited nor determined by his performance on the first job.

This principle is recognized by many supervisors and personnel men. The frequent shifting of an employee from job to job until one is found that he can perform satisfactorily indicates that the success of an employee on one job does not necessarily indicate his success on another.

Negative transfer or interference

If the individual who has taught himself the "hunt and peck" system on a typewriter decides to learn the touch system, he must give up the use of all, or most of, the habits he has already formed. It is usually as difficult to break old habits as to learn new ones, and sometimes much more difficult. Thus, a person who has thoroughly learned the wrong way to do a thing before he is taught the right way to do it has the double task of unlearning old methods and learning new ones.

Unlearning errors is a most wasteful form of learning.[32] It is not surprising, therefore, to find experiments showing that when learners have previously been taught *one* method of doing a task, they have more difficulty in learning a new method than do learners who have had no previous experience on the job at all. This principle is known as negative transfer, interference, or inhibition. Its application to the industrial situation suggests that the time to teach employees the correct way to do a job is when they are placed on the job, not after they have had a few days or weeks of experience in learning incorrect methods.

---

[32] E. L. Thorndike, E. O. Bregman, J. W. Tilton, and E. Woodyard, *Adult Learning* (The Macmillan Company, 1928), p. 183.

9

Industrial Inspection[1]

NEARLY all industrial products undergo some form of inspection. Sometimes inspection is accomplished by mechanical or other automatic devices, but often it consists of an inspection for appearance and is therefore dependent upon the visual and manipulative skills of the inspectors. The skills and capacities necessary for accurate inspection may be measured by means of certain psychological tests, and the training program that is given to new inspectors may also function more efficiently if the instruction is organized around basic psychological principles of learning. The general fundamentals of testing and training have been discussed in the preceding chapters.

The major purpose of this chapter is to describe certain procedures that have been found helpful in the selection and training of tin-plate inspectors. It is recognized that the specific conclusions reached with regard to this one type of inspection may not apply directly to inspection of other products; but the methods described, which have been found to result in definite improvement both in the selection and training of these inspectors, apply with very little modification to the problems of selecting and training employees for any inspection job. It should be emphasized that the present discussion deals with *methods*. Application of these methods to tin-plate inspection has been chosen as a means of describ-

---

[1] This chapter is based largely upon data previously published by Joseph Tiffin and H. B. Rogers, "The Selection and Training of Inspectors," *Personnel*, XVIII (1941), pp. 14–31.

ing these methods because several years of first-hand experience in their application on this job have given direct evidence of their practicability.

The job of tin-plate inspecting, ordinarily called assorting, is essentially an inspection for appearance that is made while the inspector turns the sheets of tin plate from one stack to another. As the sheet is turned, the inspector makes a decision from the appearance and feel of the sheet as to whether it is a prime or a second or contains one of a number of possible defects. The standard size sheet is approximately 28 by 30 inches, though this varies with the specific order. In the plant studied, the work is done by women assorters.

Supervisors have generally felt that inspectors on this job do not reach their maximum performance until they have had approximately six months of experience. The inspectors are paid on a straight and uniform hourly rate. At the time of these experiments approximately 300 girls were employed in this operation.

Measuring the Accuracy of Inspectors

In many production jobs it is possible to determine the relative efficiency of a group of employees from records of production or earnings. Likewise, in some inspection jobs a reinspection of samples of material is often used to check the accuracy of the original inspectors. On the inspection job of the present study, however, neither of these indications of employee efficiency was available. The inspectors were paid on a straight hourly rate and the mechanics of wrapping and shipping made it impractical to reinspect samples of the material.

It was therefore necessary to devise a method of measuring the accuracy of the present inspectors. Since this problem often arises in an inspection department, we will give a fairly detailed description of the method used. This method resulted in a criterion of job performance that was used later in the investigation.

After considerable preliminary experimentation, a coded stack of tin plate containing 150 sheets was assembled. This stack was assorted (inspected) by 150 operators. This coded stack was made up of 61 prime sheets (sheets satisfactory in every respect) and second sheets (sheets containing a minor surface blemish or uneven coating of tin), 30 sheets containing appearance defect No. 1, 26 sheets containing appearance defect No. 2, 13 sheets containing appearance defect No. 3, and 20 sheets containing a weight defect. The prime and second sheets were included in a single category for the purpose of the experiment because with repeated assortings the prime sheets tend to become scratched and thus to become seconds. The other defective plates included in the coded stack remained the same regardless of the number of times the stack was assorted.

The defects will be referred to by number rather than by name because it is the method of measuring the accuracy of the inspectors rather than their accuracy on specific defects that is of primary importance.

Each sheet included in the coded stack was carefully selected in advance and was known definitely to be either a prime or a second, or to contain one of the defects selected for study. The 150 sheets were numbered in random sequence. As the inspectors assorted the sheets, they called aloud their judgment of each sheet as to whether it was a prime or a second or contained one of the defects, and, if so, which one.

As each operator assorted the 150 sheets her judgments were recorded by an observer. She was also timed with a stop watch. This procedure resulted in data from which it was possible to determine both the speed and the accuracy of the assorters studied.

The coded stack was scored for each girl according to the total accuracy on the 150 sheets and also for the accuracy on the specific defects. This procedure resulted in five specific measurements of accuracy for each girl. It also resulted in a

measurement of the reliability of each of the five methods of scoring the coded stack test. The five resulting measurements with their respective reliabilities are summarized in Table 24.

TABLE 24

THE FIVE MEASUREMENTS OF ACCURACY YIELDED BY THE CODED STACK, WITH THE RELIABILITY OF EACH

Method of Scoring	*Reliability*
Mixed sheets	+.90
Appearance defect No. 1	+.86
Appearance defect No. 2	+.87
Appearance defect No. 3	+.68
Weight defect	+.74

The reliability figures given in Table 24 were obtained by correlating correct inspections on odd versus even items for each defect, as described on page 64. The reliability indicates the extent to which repeated or duplicate measurements of each girl by means of the coded stack test would result in the same score for her for the defects in question. If repeated measurements would give exactly the same score to all of the girls on repetition of a given test, the reliability of that test would be +1.00. If, on the other hand, the present scores given to the girls were entirely the result of chance and would, therefore, bear no relation at all to repeated measurements on the coded stack test, the reliability would be .00. It may be seen from Table 24 that the reliabilities vary from .68 to .90. While these reliabilities are not so high as might be desired, they compare favorably with the reliabilities of many other industrial criteria that have been used to measure the successfulness of employees.

As the inspectors assorted the coded stack, they were allowed to set their own speed or rate of inspection. This procedure resulted in very great differences in speed. The time required for the various girls varied from eight to forty-eight minutes. It was first thought that this uncontrolled time factor would be a marked handicap to this method of measuring inspector accuracy. After careful consideration, however, it was thought best to retain an indefinite time limit

procedure because only under such circumstances would it be possible to determine for each girl the maximum accuracy of which she was capable. In other words, the coded stack test was deliberately given under conditions that would cause every girl to bring to bear every bit of inspecting knowledge she had and that therefore would allow her to reach a level of accuracy that was close to, if not equal to, the maximum accuracy of which she was capable.

The results of the coded stack test as administered to the 150 girls are summarized in graphic form in Figure 108. The six curves shown in this figure are frequency distributions obtained from the 150 inspectors.

The time curve in the upper left-hand corner indicates the number of girls completing the 150 sheets of tin plate in each of the times specified along the base line. The most rapid assorter, it will be noted, completed the test in 8 minutes, whereas the slowest required 48 minutes. The curve is highest over the region of 14 to 21 minutes, indicating that a plurality of the girls required between 14 and 21 minutes to assort the 150 sheets under the conditions of the test. The average time required, as indicated in Figure 108, is 20.1 minutes.

In the curve immediately below this, the ordinate again represents the number of girls and the base line represents the percentage of accuracy for the entire stack of 150 sheets. The average accuracy for the 150 sheets under the conditions of this test was, as shown in Figure 108, 78.5 per cent. This does not mean that the accuracy of these assorters in a normal assorting situation is only 78.5 per cent. The assorting skill required to inspect the coded stack is greater than that required in a normal assorting situation, because more defective plates of various types are included in the test than are found in the run of the mill and the defective plates were not included in any sequence or groupings. The test as given measures differences among the girls in their skill on this job but probably gives to each girl a somewhat lower score, that

is, a poorer accuracy, than she customarily reaches in routine daily assorting.

The remaining four curves in Figure 108 give similar distributions for the four specific defects studied. Figure 108 shows, among other things, that for the four defects studied the average inspector is most accurate on appearance defect No. 2 (75.5 per cent), and least accurate on appearance defect

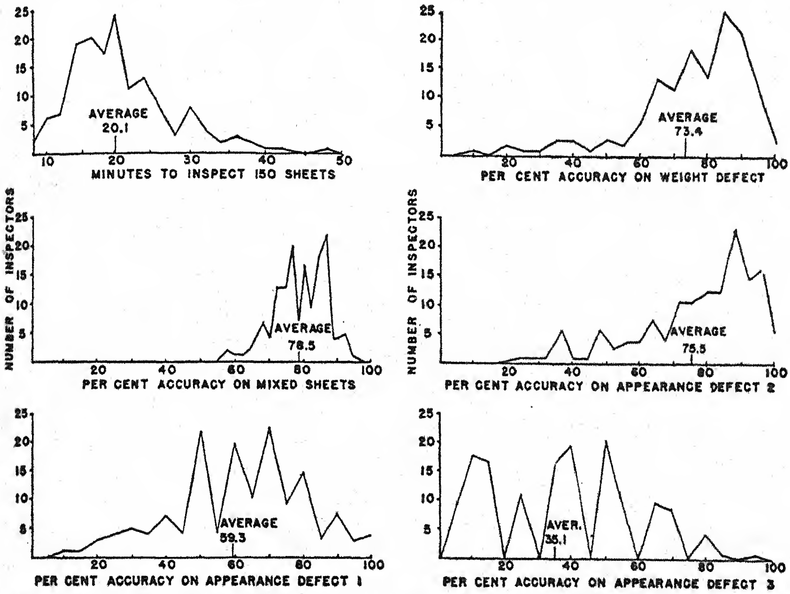


FIG. 108—Distributions of results on a codes stack test of 150 tin-plate inspectors.

No. 3 (35.1 per cent). The data also show a great variation among the various girls in their accuracy on any specific type of defect. For example, 8 girls out of the 150 detected only 5 per cent of appearance defect No. 3, whereas one girl spotted 95 per cent of the sheets containing this defect. There is nearly as wide a variation in accuracy for the other four defects studied.

Inspection of Figure 108 also reveals that the amount of variation among the girls, that is, the range from the poorest to the best, is smaller for the curve representing total accuracy

than for the several curves representing accuracy on specific defects. The reason for this is that only a slight relationship exists between accuracy in detecting any one defect and accuracy in detecting other defects. Thus, if a girl is far above average in detecting one defect she is likely to be at average, or possibly below average, in detecting other defects. This fact tends to give scores for the total 150 sheets of mixed plate more nearly alike than are the scores for the various individual defects.

It will be noticed that the average accuracy on mixed plate is higher than the accuracy on any specific defect. This difference is due to the fact that a majority of the sheets included in the coded stack were primes or seconds and that the acceptable sheets were seldom called defective, whereas the defective sheets were frequently called primes or seconds. In other words, there was a tendency to overlook the defects. Hence, the accuracy for the total stack, which included a majority of primes and seconds, and a minority of defective sheets, was found to be higher than the accuracy on any specific defect studied.

The next question investigated was the relation between speed and accuracy in detecting the different types of defect. The results for this part of the investigation are summarized in graphic form by the curves in Figure 109. Each of the curves in this figure is a logarithmic curve of best fit[2] representing the relationship between speed and accuracy. The equations of the curves shown in Figure 109 are as follows:

Percentage accuracy on appearance defect 1 $= 24.0 + 33.2 \log_{10} (\text{Time}-7.5)$
Percentage accuracy on appearance defect 2 $= 55.5 + 19.5 \log_{10} (\text{Time}-7.5)$
Percentage accuracy on appearance defect 3 $= 9.4 + 28.4 \log_{10} (\text{Time}-7.5)$
Percentage accuracy on weight defects $= 72.7 + 5.6 \log_{10} (\text{Time}-7.5)$

In using these equations, it should be remembered that the word "Time" refers to the number of minutes used in assorting 150 sheets. It is understood that the percentages

[2] The curves were fitted according to the mathematical criterion of least squares.

of accuracy that the equations yield are for the conditions under which the coded stack test was given.

It is interesting to note that in detecting off-weight sheets the accuracy is about the same whether the test is completed in 10 minutes or in 40 minutes. In other words, accuracy in detecting off-weight sheets for the 150 girls tested is not

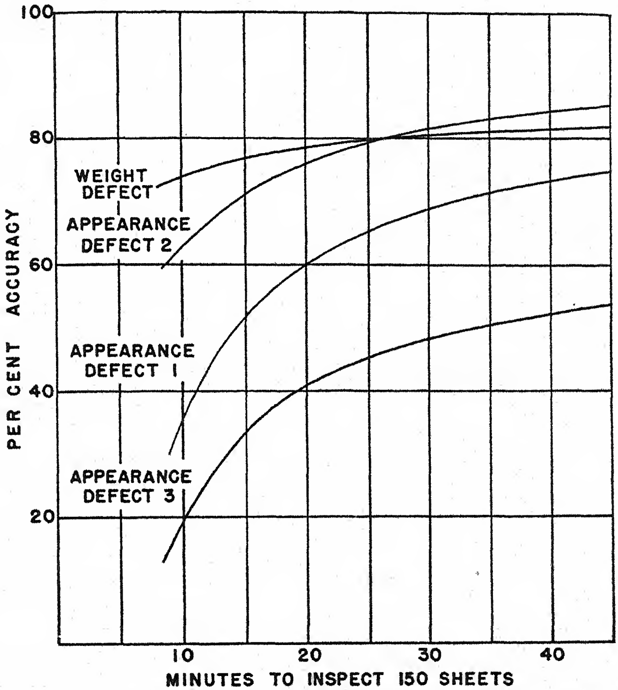


FIG. 109—Relationships between speed of inspection and accuracy in detecting four defects.

appreciably affected by speed of inspecting within the time limits found for these girls. To some extent the same general situation is true for appearance defect 2, although in detecting this defect we note a beginning tendency for accuracy to decrease with increased speed. This tendency is still more pronounced in the case of appearance defects 1 and 3. For example, we find that when the stack is assorted in 10 minutes,

only 20 per cent of the sheets containing appearance defect No. 3 are spotted and only 35 per cent of those containing appearance defect No. 1.

On many inspection jobs, certain defects may be spotted equally well at either fast or slow rates of inspection, but other defects cannot be spotted with satisfactory accuracy above a certain inspection speed. These curves clearly reveal this situation for this inspection job. Defects of weight are noted almost as well when the 150 pieces are inspected in 10 minutes as they are when 40 minutes are used to inspect this number of sheets. It is important for the head of an inspection department to know how accuracy is affected by varying speed. When he possesses this information he is more able to set the speed of inspection so as to achieve satisfactory quality without wasting time in overcareful inspection.

The coded stack used in this study furnished not only a means of measuring the accuracy of the inspectors but also data from which the relationships between speed and accuracy plotted in Figure 109 were determined.

The Determination of Tests for the Placement of Inspectors

The tests selected for tryout consisted of a series of vision tests,[3] the Minnesota Rate of Manipulation Test,[4] the Purdue Hand Precision Test,[5] and, in addition, measurements of reaction time, strength of grip, height, weight, and age.

Table 25 summarizes the interrelations of the several tests in terms of the correlations[6] between them. According to the technique used for computing these correlations, the maximum value the figure theoretically could reach is + 1.00. This value would be reached in the case of a perfect positive relationship between the two variables under consideration.

---

[3] See page 185.

[4] See Table 13 on page 133.

[5] See page 128.

[6] The correlations in Table 25 are product moment correlations, computed as described in Appendix A, page 511.

The minimum figure the correlation could reach is −1.00, which would represent a perfect negative relationship.

The first figure mentioned in the table, it will be noted, is −.49, which indicates the relation between speed and accuracy on mixed plate. This figure means that there is a tendency for those who are above average in speed to be below average in accuracy, and vice versa. It will be noted further that the correlations between speed and accuracy for the various specific defects are all negative and that the amount of negative relation between speed and accuracy on off-weight plate is the smallest found in the group. The intercorrelations between accuracy for the various specific defects are all small, the largest being only +.35. This finding substantiates the statement previously made that there is no decided tendency for girls who are above average in detecting one

TABLE 25
Correlations of Test Results of 150 Tin Plate Inspectors

	Speed	Mixed sheets	Appearance defect 1	Appearance defect 2	Appearance defect 3	Weight defect
Mixed sheets	−.49					
Appearance defect 1	−.42	+.80				
Appearance defect 2	−.49	+.68	+.24			
Appearance defect 3	−.36	+.58	+.35	+.21		
Weight defect	−.05	+.29	+.22	+.02	+.12	
Visual acuity (distance)	−.26	+.22	+.28	+.16	+.19	+.02
Visual acuity (near)	.00	+.09	+.08	+.13	+.15	−.10
Vertical balance of the eyes	−.43	+.03	+.25	−.13	+.08	+.01
Height	−.07	+.21	+.05	+.02	+.10	+.13
Weight	−.20	+.25	+.09	+.14	+.04	−.05
Age	−.11	−.16	−.09	+.03	−.12	−.21
Experience	+.05	−.05	−.07	.00	+.06	−.16
Strength of grip	+.02	+.17	−.05	+.19	−.07	+.14
Minnesota Rate of Manipulations	+.01	+.06	−.10	−.03	+.03	+.14
Speed of Reaction	−.02	+.12	+.05	+.10	−.07	+.03
Purdue Hand Precision	+.05	−.13	−.06	+.12	+.01	+.34

type of defect to be above average in detecting other types of defects.

The general conclusion that may be drawn from the correlations between the vision tests and the specific ways of scoring the coded stack is that the vision tests clearly tend to pick the girls who are the most accurate on the job. Although two correlations are slightly negative, the remaining thirteen are positive. The distance visual acuity test gave the highest positive correlations with accuracy in detecting all three appearance defects.

Table 25 also shows correlations between accuracy, as revealed by the coded stack test, and the height, weight, age, and experience of the operators. It will be noted that the taller girls tend slightly to be somewhat slower than the shorter girls (correlation between height and speed is $-.07$). Taller girls, however, tend to be more accurate on the several phases of the coded stack test, as indicated by the fact that the remaining correlations in this row are all positive. Likewise, in considering body weight we note that the heavier girls are somewhat slower but more accurate on everything but the off-weight sheets. The older girls, however, are not only slower but less accurate on everything but appearance defect 2. The more experienced girls are slightly faster but less accurate on nearly every defect.

Correlations between the remaining tests and the various ways of scoring the coded stack are given in the last four rows of Table 25. While the correlations shown in the remaining rows are somewhat inconsistent, at least one value is high enough to indicate a definite and real relationship. This value is the correlation of $+.34$ between the Purdue Hand Precision Test and accuracy in detecting off-weight plate. The hand precision test is recommended, therefore, for inclusion in a battery of employee placement tests as discussed in the last section of this chapter.

The student may wonder why correlations have been included in Table 25 that are so low that no appreciable

relationships are indicated. The reason for including these low correlations in this table is to emphasize again the fact that in setting up any battery of employee tests it is always necessary to start with more tests than one expects finally to retain. Many of one's guesses are likely to be wrong. In the present experiments, the tests that were found of no value were strength of grip, Minnesota Rate of Manipulations, and reaction time. These represent the bad guesses in finding aptitude tests for the job. However, it should not be assumed that these tests are generally of no value, for they might be the very ones to "come through" best on some other job.

The following general conclusions from the test results will give some indication of the kind of results such a procedure may yield:

1. By means of a coded stack of material, it is possible to measure the accuracy of inspectors.

2. These measurements reveal whether any relationship exists between speed and accuracy of inspection, and, if so, how much relationship.

3. Results of this procedure also reveal whether a good inspector is accurate on all types of defect or more accurate on some defects than on others. Such information is of value in training or re-training inspectors.

4. The procedure also reveals which defects are most difficult for inspectors to detect. This information is also of value in a training program.

5. A battery of vision tests was found to be of importance in placing persons on this job of inspection for appearance. It is probable that the same tests would be of value in placing employees on any other job of inspection for appearance.

A Battery of Placement Tests for the Inspectors

A study of the correlations shown in Table 25 resulted in the setting up of four qualifications that a girl must meet

before being trained as a tin-plate inspector. These requirements were:

1. Pass the near and far visual acuity tests and the vertical balance test.
2. Be at least 5 feet 2 inches tall.
3. Weigh at least 118 pounds.
4. Score not over 2.00 on the Purdue Hand Precision Test. (See page 128.)

An indication of the effect of placing inspectors upon the basis of these recommendations may be seen from Table 26, which is based upon data gathered in this investigation.

TABLE 26

A Comparison of the Job Performance of 150 Inspectors Selected at Random with the Job Performance of 28 Inspectors Who Met Four Basic Requirements

	Average of all 150 inspectors	*Average of 28 inspectors who met qualifications on four tests*	*Difference in favor of inspectors passing tests*
Time to assort 150 sheets	20.1 min.	21.1 min.	5% slower
Accuracy on mixed sheets	78.5%	81 7%	4.1% more accurate
Accuracy on appearance defect 1	59.3%	68.4%	15.3% more accurate
Accuracy on appearance defect 2	75.5%	83.1%	10.1% more accurate
Accuracy on appearance defect 3	35.1%	38.0%	8.3% more accurate
Accuracy on weight defect	73.4%	76.4%	4.1% more accurate

The inspectors meeting the qualifications, though 5 per cent slower on the coded stack, were from 4 per cent to 15 per cent more accurate in detecting the several defects. It might be suspected that, since the inspectors in the qualifying group were slower by 5 per cent on the coded stack than the average of all inspectors, the greater accuracy of this group is simply a reflection of this slower speed. However, reference to the equations showing the relation between speed and accuracy (see page 302) and the curves of these equations, plotted in Figure 109, shows that the speed differential would account for a maximum of only 1.8 per cent difference in accuracy. This result may be found by solving the equation for percentage of accuracy on appearance defect, using a time value of

20.1 minutes, and then solving this equation again using a time value of 21.1. The solutions will yield accuracy percentages, respectively, of 60.5 per cent and 61.6 per cent. The difference between these two percentages, 1.1 per cent, divided by the percentage 60.5 per cent, gives a percentage difference of 1.8 per cent. This difference is considerably less than the difference in accuracy on any one of the defects. It seems conclusive, therefore, that the greater accuracy of the qualifying group is due primarily to the visual and other test characteristics of this group, rather than to the fact that they inspected at a slightly slower speed than the average of all inspectors studied.

Relation between Psychological and Motion Study Analysis of the Job

The discussion thus far has dealt chiefly with the statistical analysis of the data obtained from supervised tests on the coded stack and from various psychological tests. It seemed desirable to check these results with records of performance under actual shop conditions, and for this purpose micromotion studies of representative operators were made.

Motion pictures were taken, at 1,000 frames per minute, of twelve inspectors selected from those who had taken the coded stack test. In this group were two who had been rated fast and accurate, two rated fast but inaccurate, two rated slow and accurate, two rated slow and inaccurate, and four rated average in both speed and accuracy according to the test stack data. The pictures were taken in the shop under normal working conditions, and because of the general noise and shop activity, the operators were not aware of the exact moment at which the pictures were taken.

Obviously the pictures did not reveal the accuracy with which the inspector detected the defects. However, they did show the normal speed of the operator, the relative body activity or exertion, the various methods used in grasping and moving the sheets, and the disturbance in the rhythm whenever a defective sheet was found. In general the speeds

indicated by the analysis of these films correlated fairly well with the speeds indicated by the test stack.

The pictures were analyzed in an attempt to detect individual motion characteristics that might indicate causes for differences in accuracy. Since the detection of all but one class of defects depends upon vision, the eye fixations were plotted in relation to the positions of the hands and the positions of the sheet.

In the normal handling of the sheets the several stacks of sheets that have not been inspected, the sheets of first quality, those of second quality, and those containing various types of defects are aligned along a bench. The operator picks up a new sheet from the pile at her left and turns it over toward the right onto the adjacent pile of first-quality sheets where it normally remains. If a defect is discovered, the sheet is moved to the proper pile farther along the bench.

For most of the operators it was found that the eyes tend to follow the movements of the hands. Thus as the hands move toward the left to grasp a new sheet, the eyes attempt to scan the full top surface of the sheet within the very short period of time required for the grasp. As the sheet is picked up and turned over onto the adjacent stack, the eyes tend to follow the movement of the sheet which, in the upright position, presents only one edge toward the eyes. As the sheet approaches the right-hand stack, the eyes attempt to scan completely the second surface of the sheet; but as the release period is very short in duration, the scanning period is likewise very brief.

In this sequence, the eyes are attempting to see an object that is almost continually in motion, and it was found that the operators who most nearly followed this pattern had the lowest accuracy ratings in the test-stack data. This result is logical since it is more difficult to see a moving object and the continual attempt to do so reduces the accuracy of the operator.

On the other hand, the accuracy ratings were higher for those inspectors whose eye movements were more nearly

opposite the sequence described above. For example, the inspector mentioned above, who was both fast and accurate according to the films and actual production records, used the following cycle of eye movements: As the hands moved to the left to pick up a new sheet, the eyes were scanning the second side of the previous sheet. The fact that this previous sheet was not in motion contributed to the ease of seeing possible defects, and the scanning period was increased to include nearly the full handling time of the next sheet. When the next sheet approached the second pile to the right of the operator, the eyes quickly shifted toward the left, to the first side of the top sheet on the new pile where the scanning period included the time for disposal of the previous sheet (see Figure 110).

In this second cycle of eye movements, the eyes move quickly from one motionless object to the other and thus gain greatly increased scanning periods for detecting the defects. Accuracy data both from the test stacks and from production records correlated very closely with the eye-movement patterns of the twelve inspectors thus analyzed.

It has been demonstrated experimentally that the eyes can see a motionless object more accurately than one that is in motion, and the longer the scanning period, the greater the opportunity to observe defects. If the increased scanning periods are longer than necessary for effective inspection, the operator tends to handle the sheets faster. It is logical, therefore, that the girl who, accidentally or otherwise, has acquired this improved routine shall be both fast and accurate.

This cycle is opposite the more natural cycle of following the hands with the eyes; but as in many other operations, the natural way is not necessarily the best way. This less natural, but more effective, co-ordination of the eyes and hands is simple to learn, but because it is less natural, its technique must be explained and taught to the operator along with the other requirements of the job. This technique might be classed as one of the "tricks of the trade" not readily discovered without adequate instruction.

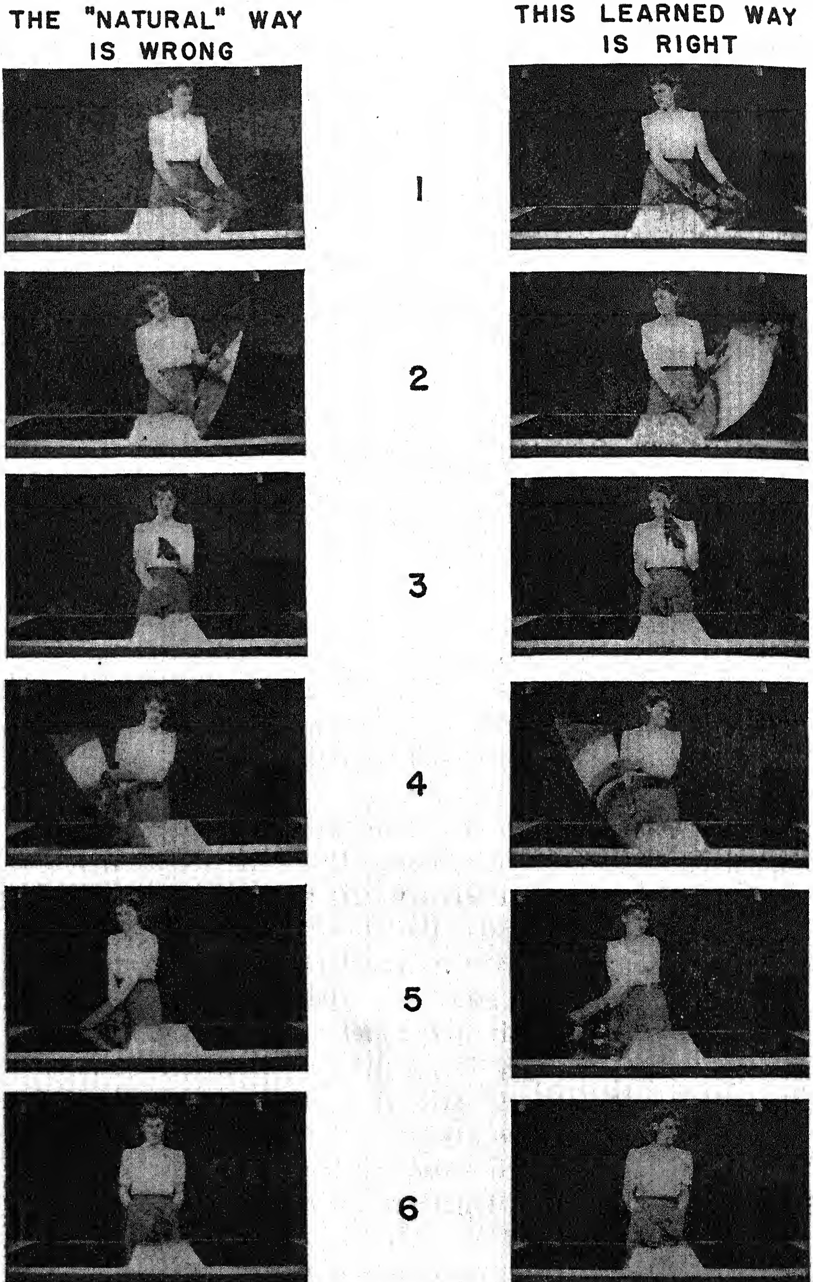


FIG. 110—Sequence of sheet movements and eye-movement patterns in inspect-

The Training of Inspectors

Several specific needs for training were revealed by the experimentation described above. In the first place, the accuracy of inspectors on the job was found to be far from the level of satisfactory performance that management considered desirable. Reference to Figure 108 shows that, in the case of some of the defects, an appreciable number of the present employees were unable to spot over 10 or 20 per cent of the defective material. The relationship discovered between accuracy and visual capability resulted first of all in the recommendation that those girls low in accuracy on defects of appearance immediately see an eye specialist. Subsequent study of the inspectors in the plant revealed that those girls who had received professional eye care that improved their performance on the vision tests were doing a more satisfactory job of inspection than those girls who still were unable to pass all of the vision tests.

A second effort made to upgrade the performance of employees on the job consisted of a series of inspectors' conferences. The inspectors were brought into conference in groups of approximately 20, on company time. The first thing done in each of these conferences was to pass out a slip to each girl on which was printed her accuracy on the coded stack as a whole and on each of the four specific defects studied. This information gave each girl a clear indication of her weak as well as her strong points and resulted, in most cases, in a definite desire on her part to improve her level of performance. The fact that little correlation was revealed between accuracy on the different kinds of defects resulted in the fortunate circumstance that a girl who was very low in accuracy on one defect, or possibly even two defects, was likely to be average, or even above average, in accuracy on

ing tin plate. In the "Natural" way the eyes follow moving objects. The inspector thus sees the full sheet only in steps 1 and 5; she tries to see it while it is in motion in steps 2 and 4; she sees only the edge in step 3; and merely watches her hands in step 6. In the Right Way, which must be learned, the inspector sees a motionless sheet during all 6 steps. This way nearly doubles the available inspection time.

the other defects. This fact, coupled with the further fact that most of the girls were low in accuracy on at least one defect, prevented any employee from feeling that she had been singled out because of her low-quality work. Each inspector attended two of these training conferences, which lasted an hour and a half each and were separated by a period of two weeks. The time was spent in thorough discussion and demonstration of the various kinds of defects found in tin plate.[7] The demonstration also included a review of the micromotion films and an explanation of differences between the patterns of eye movements with respect to the moving sheet. The films illustrated these eye-movement patterns and provided a means of teaching the most effective sequence of visual fixations. General improvement in performance following these conferences was noted.

The second innovation in training procedures that directly resulted from this investigation dealt with the training of new employees. The procedure followed in the past consisted in placing a new girl between two experienced inspectors. The expectation was that new employees would be able to ask questions of the more experienced employees. The actual results indicated that this procedure had not resulted in adequate training of new inspectors. Since hindsight is often better than foresight, it was easy to conclude after the investigation that experienced operators do not wish to be bothered with stopping frequently to answer the questions of a new girl. This knowledge resulted in each new employee finally setting up her own methods and standards of performance. It is not surprising that under such circumstances the standards varied markedly among the employees on the job.

The new procedure consisted in assigning an experienced and capable inspector as an instructor to each group of four new employees. The old employee was taken off the job of

[7] This procedure has been suggested by J. H. Mitchell, "Subjective Standards in Inspection for Appearance," *The Human Factor*, London, IX (1935), pp. 235–239.

inspection, and since she had no assigned inspection work to perform she was able to devote all of her time to instruction of the new employees. This procedure resulted not only in much more uniform standards of inspection among the newer employees than had heretofore existed but also in a marked reduction of the time necessary for bringing a new girl up to a satisfactory level of performance. As the newer girl became more adjusted to the job, less and less time of the instructor was called for; and after a certain level of efficiency was reached, the instructor automatically resumed her former work. Under this system the older employees assigned as instructors took considerable pride in the fact that they had been selected for the job of breaking in the newer inspectors, and the method was favorably received both by the older employees and the newer employees, as well as by the management.

The general conclusions from this work are that if adequate inspectors are to be obtained they must be selected with scientific precision and trained with corresponding thoroughness. No amount of desire on the part of a new employee to be a good inspector will result in satisfactory job performance if the employee does not have the visual qualifications, the necessary eye-hand co-ordination, or the physical stamina that experiments show to be necessary for the job. And in like manner, no amount of qualification for the job will result in satisfactory performance when a systematic training program covering standards and procedures as well as "tricks of the trade" is lacking.

Inspection with Precision Instruments

A further illustration of the application of psychological methods in the field of inspection may be found in a study of precision instrument inspection.[8]

Modern industrial production is becoming more and more

---

[8] C. H. Lawshe and Joseph Tiffin, "The Accuracy of Precision Instrument Measurement in Industrial Inspection," *Journal of Applied Psychology*, XXIX (1945), pp. 413–419.

dependent upon the accuracy of this type of inspection. Virtually every precision instrument calls upon the operator to exercise judgment in determining proper "feel," "tension," "drag," or other characteristics, yet, in spite of all that is known about the variability of human judgments, little attention has been given to the importance of such variability as it may affect the accuracy of precision instrument measurement. Recently an investigation was conducted to examine the accuracy and variability of employee measurements made with certain precision instruments. Data were collected in two different plants. The first of these was engaged in the manufacture of variable pitch propellers for aircraft, the second, in the manufacture of precision parts for aircraft and automobile engines.

Approximately two hundred people are employed in the inspection department of the first plant. Their jobs were analyzed by job classifications in order to determine what precision instruments are used and what tolerances are demanded on each job. Frequency counts were then made to determine which instruments or combinations of instruments are used in the largest number of classifications and by the largest number of employees. On the basis of this count, twenty instruments and combinations were chosen as being most important in this particular plant.

A dimensional control laboratory

A room was set aside as a dimensional control laboratory and twenty booths or inspection stations were set up. Each booth was numbered and in it were placed one of the twenty instruments, a standard part from the plant, and a simplified working drawing which indicated one dimension to be measured with the instrument provided. When an employee entered the room, the attendant determined his job classification and provided him with an appropriate work-sheet for each of the stations containing work samples from his job. Each employee was tested on only those instruments which he

uses on his particular job. He was encouraged to make five measurements and then to record his best judgment as to the dimension. The readings thus obtained were compared with so-called "true" dimensions, which were determined by means of ultra-precision instruments in combination with Johansen blocks. Instruments utilized in the performance testing were checked and adjusted periodically to insure constancy.

Emphasis of testing

This performance testing procedure was organized in connection with a training program, and its primary function was to identify persons in need of training. Plans for a maintenance program were also made with provisions for re-testing employees every three months. It was also planned to utilize the laboratory to supplement seniority in determining adequacy in connection with transfers and promotions. The program was instituted with the knowledge and backing of line supervision and of the union in the plant.

Results obtained at eleven of the twenty stations are presented in Figure 111. The particular stations selected for illustration were chosen in terms of general familiarity with the instruments used and not because of any peculiarity in the findings; they are truly representative.

In Figure 111, the open bars represent the percentage of inspectors tested who obtained readings within the established tolerance. The solid bars represent the percentage of persons tested who failed to meet the standard. As already stated, not all the inspectors were tested at each station; instead, the sample contains only those who use the instruments on their jobs. This accounts for the fact that the N's range from 117 to 162. The figure indicates that the percentage of inspectors meeting the various standards ranged from a high of 66 per cent on the inside micrometer to a low of 9 per cent on the inside caliper in combination with the six-inch micrometer. The pattern of performance on the various vernier micrometers also seems significant. It will be noted that 43 per cent

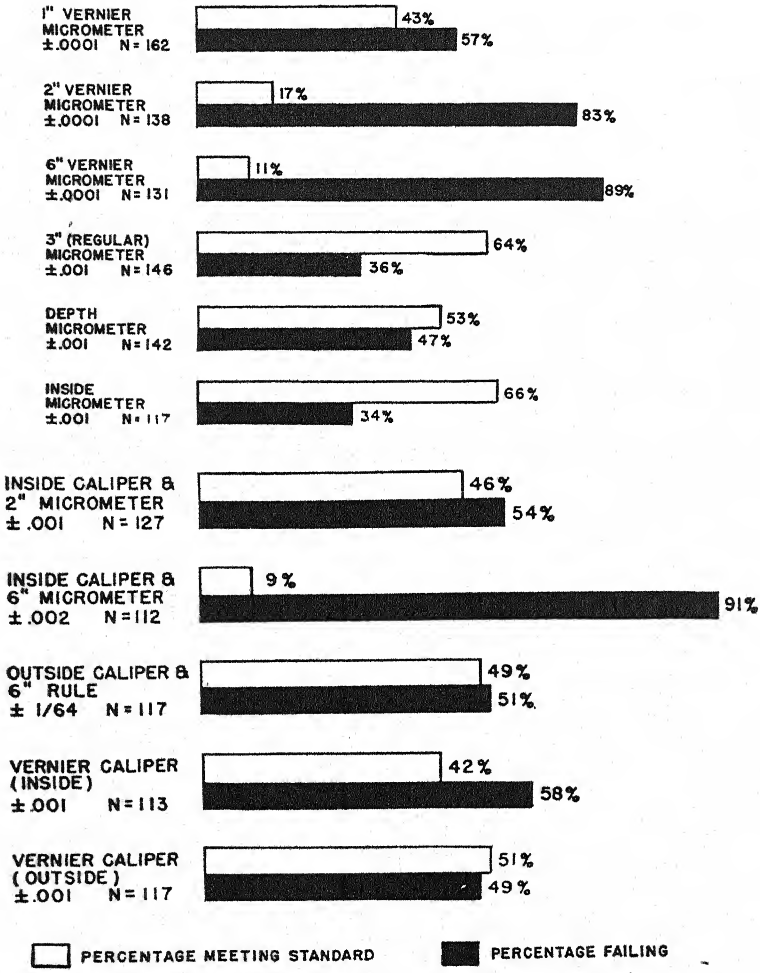


FIG. 111—The percentage of inspectors passing and failing various precision measuring instrument performance tests in an aircraft propeller plant. The open bars indicate the percentage meeting the standard and the solid bars indicate the percentage failing.

of those tested met the standard with the one-inch micrometer, 17 per cent with the two-inch, and only 11 per cent with the six-inch. The varying tolerances established for the instruments are the same as the tolerances which job analyses

indicated had been established by the engineering department and are identical with those encountered in the shop.

After the completion of the study summarized, a similar investigation was conducted of forty-five toolmakers from another plant. The percentages of toolmakers who were able to measure various parts within specified tolerances were similar to those cited above for the inspectors. The results further showed that the larger the part, the less accurate the measurement. In both plants where these studies were conducted, training programs in the use of precision instruments were found to be needed, and were instituted after the surveys were completed. These studies typify the manner in which psychological measurement of employee performance markedly helps in identifying areas where operators are insufficiently trained, or for other reasons are not performing their jobs in an entirely satisfactory manner.

10

Industrial Merit Rating

MERIT rating is a systematic evaluation of an employee by his supervisor or by some other qualified person who is familiar with the employee's performance on the job. Merit ratings are usually made by means of a standardized form that is adapted to the needs of the particular industry. Usually these ratings are made at periodic intervals. A merit rating thus becomes a permanent part of an employee's record with a given company, and, at least in theory, is a part of the record that may be used by management in subsequent promotion, transfer, or layoff.[1]

The Growth of Merit Ratings

A survey conducted in 1939,[2] which covered 64 companies employing from 500 to more than 100,000 employees, revealed that merit rating in some form was then in use by approximately one third of the plants surveyed. Since that time, many factors have tended to increase still further the number of industries using a merit-rating system. Contracts that have been signed by management with labor organizations frequently contain a clause to the effect that an employee has a right to his job at the time of layoff in proportion to the amount of the employee's seniority, if other things such as

---

[1] An excellent summary of practice and experience in industrial merit ratings appears in "Employee Rating: Methods of Appraising Ability, Efficiency, and Potentialities," *Studies in Personnel Policy*, No. XXXIX (National Industrial Conference Board, New York, 1942).

[2] R. B. Starr and R. J. Greenly, "Merit Rating Survey Findings," *Personnel Journal*, XVII (1939), pp. 378–384.

"job performance, skill, and ability" are equal. Such a clause in a labor contract necessarily involves some method of measuring or evaluating whether the "other things" are equal. Merit ratings have frequently been proposed as one means of meeting this situation. Whether, eventually, merit ratings or some other tool or technique of management will be adopted to meet this need, at present an ever-growing number of industries are making a serious attempt to solve the problem with merit ratings. Whatever may be the industrial future of this technique, and however keenly one may be aware of the criticisms that may be leveled against many systems of rating now in use, it seems clear that merit rating is likely to remain a sufficiently important part of personnel administration for some time to come to justify our giving rather careful consideration to its discussion.

One point should be made clear immediately: while the term "merit rating" is new, the rating of men by supervisors is by no means new. Such rating has been carried on as long as industry has been in anything like its present mass-production form. Supervisors have *always* rated men, and it is no doubt true that the ratings made in random, slipshod, and unsystematic fashion, unrecorded and undefended, have in the past been just as important, if not more important, in determining whether a given employee should hold his job as any rating made by means of a modern merit-rating chart. The changes that came with systematic merit rating, then, were not changes that involved a making of ratings where none had existed before; rather they were changes that involved the transfer of ratings from haphazard, random, and frequently irresponsible judgments of supervisors made perhaps during the heat of a quarrel, to ratings made calmly, deliberately, systematically, and in a manner that made the ratings, if not completely comparable from one employee to another, at least much more comparable than were the older, haphazard evaluations of employees by supervisors. The question, then, is not whether supervisors should rate their

employees—this always has been and probably always will continue to be done—but whether the use of a formal merit-rating system is likely to increase the value of such ratings both to management and employees.

Purposes of Merit Rating

Among the more common purposes that merit rating is intended to serve are the following:

1. **Promotion.** The identification of men who are eligible for promotion either to a higher-rated or to a supervisory job. Union contracts often contain a clause providing that seniority shall govern promotions only when the "ability, skill, and job performance" of applicants are equal. Management has frequently proposed the use of merit ratings to indicate whether or not these other factors are equal. Unions, however, have been very reluctant to accept merit ratings as a satisfactory measure of "ability and skills." This reluctance has been apparent not only in cases where the union-management relationship is characterized by frequent disagreement on other topics, but also, in almost the same degree, even in cases where the union and the management are in close harmony on many other subjects. This very fact, namely, that it has been extremely difficult to get merit rating accepted even by unions that co-operated with management to a considerable degree on many other subjects, such as job evaluation, suggests the possibility that many merit-rating systems may be subject to legitimate criticisms.

2. **Transfer.** The identification of men who have the necessary skills, abilities, and/or adaptability to fill a vacancy on another job or in another department.

3. **Employee improvement.** Analysis of the strong and weak points of an employee so that both he and management can properly direct their efforts toward the development of the personal characteristics, skills, or information that will make him a better employee and increase his chances of further promotion.

4. **Testing personnel tests.** The identification of groups of better or poorer employees for the purpose of testing or trying out personnel tests. (See page 55.)

5. **Layoff.** A fifth possible purpose to which considerable importance was attached when merit rating first came to the attention of management but which is not often mentioned today is the use of merit rating to identify men for layoff when production is

curtailed. Most union-management contracts today provide for strict adherence to seniority in layoffs.

Merit-Rating Systems

Several types of merit-rating systems have been developed. These systems offer various advantages and disadvantages, depending upon the primary purpose the ratings are intended to serve.

The chart system

Some form of chart system is the most widely used kind of merit-rating plan currently in use. The basic principle of the chart system is the rating of the employee on each of a number of traits or worker characteristics. The charts used differ so much from one plant to another that it would be quite unwise or even impossible to propose any single form that would be adapted to the needs of all industrial organization.[3] The wide variety both in traits rated and in number of traits rated may be seen from an analysis of the merit-rating charts of 18 companies[4] as shown in Table 27. In this table the companies are referred to anonymously by number across the top row of the table. The number of traits included varies from 21 in the case of Company 1, which is analyzed in the first column, down to four traits in the case of the companies analyzed in the last two columns. The median number of traits used is ten; therefore, if any virtue lies in considering the practice of the typical or average industry as indicating the correct practice to follow, ten traits would be the proper number to list in a merit-rating chart. Probably, however, little virtue lies in attempting to fit a merit-rating chart to the average of a typical industry; it

---

[3] A summary of the steps that should be taken in developing a merit-rating system has been published by J. E. Zerga. See "Developing an Industrial Merit-Rating Scale," *Journal of Applied Psychology*, XXVII (1943), pp. 190–195.

[4] Starr and Greenly, *op. cit.* This article refers to a survey of 16 companies, but the tabulated results show an analysis of 18 merit-rating charts.

TABLE 27

THIRTY-FIVE RATING ITEMS USED BY EIGHTEEN COMPANIES[5]

A check indicates that the item appears on the company's merit-rating chart. A number following the check indicates the weight assigned to the item. If no numbers appear it is intended to weight all items equally.

	NUMBER OF COMPANY																		
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	FREQUENCY
Quality		√8	√	√	√12		√50	√20				√	√	√	√10		√	√50	12
Quantity	√	√6	√	√	√10		√50	√25				√	√	√			√		11
Co-operation	√	√6	√	√			√50	√ 8	√50	√			√	√	√10				11
Initiative	√	√6	√	√	√10	√	√50	√ 6	√50	√	√								11
Dependability	√	√6	√			√		√ 9	√50	√				√	√10		√		10
Personality	√			√		√	√50		√50	√	√		√		√10	√			10
Health	√	√4	√	√	√ 4	√	√50		√50			√				√			10
Safety	√	√6	√		√ 8		√50	√ 7	√50				√	√					9
Industry		√6	√	√	√10	√	√50			√	√				√10				9
Versatility		√6	√	√		√	√50	√15					√				√		8
Leadership	√	√6	√	√	√10			√ 5	√50		√								8
Judgment	√	√4		√			√50		√50	√					√10				7
Intelligence	√	√6	√		√12					√	√					√			7
Attendance		√4	√		√ 4			√ 5					√					√10	6
Knowledge of job			√	√			√50			√	√	√							6
Potentiality	√			√						√	√			√					5
Habits		√4			√ 6	√										√		√10	5
Years of service					√			√15				√						√30	4
Loyalty	√	√4				√													3
Ability to plan	√					√			√50										3
Enthusiasm		√4				√				√									3
Trade skill		√6			√10		√50												3
Technical knowledge		√4	√	√															3
Dependents								√10				√				√			3
Punctuality		√4	√																2
General rating	√								√50										2
Tact	√					√													2
Suggestiveness	√			√															2
Knowledge of costs	√				√ 4														2
Fairness	√																		1
Knowledge of product	√																		1
Knowledge of equipment	√																		1
Knowledge of company policies	√																		1
Appearance						√													1
Place of residence												√							1
Number of rating items	21	19	15	14	13	12	12	11	10	10	7	7	7	6	6	5	4	4	
Method of rating (C = committee; S = supervisor)	S	—	C	S	S	S	C	C	C	—	S	—	S	S	S	—	S	—	
Explanatory phrases used on rating scale	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No	Yes	Yes	No	No	No	

[5] From Starr and Greenly, *op. cit.*

A check indicates that the item appears on the company's merit-rating chart. A number following the check indicates the weight assigned to the item. If no numbers appear it is intended to weight all items equally.

would, then, be better in any specific instance to adapt the chart to the industry intending to use it than to make it conform to a given set of characteristics, however representative this set of characteristics might be of industry in general.

There is some reason to believe that with the increasing use of merit-rating charts the tendency will be to rate fewer traits rather than more. The primary reason for this belief is to be found in the existence of the "halo" effect, a phenomenon almost always found in psychological ratings. This effect will be discussed in detail on page 333. One typical chart illustrating the chart method of merit rating is reproduced in Figure 112.

Rank-order system

When a chart system is used, there is frequently a tendency for the raters to pile up the ratings at one end of the scales, frequently at the higher end, owing to an inclination toward leniency on their part. To the extent that the ratings are concentrated at one end of the scale, the results are of little value, because these ratings will not differentiate among men who are located at the same point on the scale.

This difficulty can be avoided if a rank-order system of rating is used. Under this plan, each supervisor arranges his men in rank order, and each man's rating is then determined by his position in rank. If a system involving several traits is used, the rankings should be made separately for each trait. The actual process of ranking is usually facilitated by the use of small cards containing the names of men to be ranked. The cards may be arranged and rearranged by the supervisor as he ranks the men on each of the traits considered. The method of paired comparisons, i.e., pairing each man with every other one, is usually advantageous if not over 15 to 20 men are to be ranked.

Forced distribution system

When a department is large—consisting, say, of forty or

PROGRESS RECORD

Name__________ Dept.__________ Div.__________ Date__________

Employee's Position____________________ Job Class.__________

Note This rating will represent in a systematic way your appraisal of the employee in terms of his ACTUAL PERFORMANCE ON HIS PRESENT JOB In the interests of furthering careful analysis, the following suggestions are offered regarding the use of this form

1. Consider only one factor at a time
2. Study each factor and the specifications for each grade
3. Review upon completion to see that the rating of each factor applies exclusively to the individual's ACTUAL PERFORMANCE ON HIS PRESENT JOB
4. Comment fully at bottom of page and on reverse side upon any matter which in your opinion needs explanation

PERFORMANCE FACTORS	PERFORMANCE GRADE				
	Far Exceeds requirements of this job	Exceeds requirements of this job	Meets requirements of this job	Partially Meets requirements of this job	Does Not Meet requirements of this job
QUALITY OF WORK Accuracy Economy of Materials Economy of Time (his own and others) Neatness Thoroughness	Consistently superior ☐	Sometimes superior ☐	Consistently satisfactory ☐	Usually acceptable ☐	Consistently unsatisfactory ☐
QUANTITY OF WORK Productive Output	Consistently exceeds requirements ☐	Frequently exceeds requirements ☐	Meets requirements ☐	Frequently below requirements ☐	Consistently below requirements ☐
DEPENDABILITY Follows Instructions Judgment Punctuality and Attendance Safety Habits	Consistently dependable ☐	Dependable in most respects ☐	Ordinarily dependable ☐	Frequently undependable ☐	Consistently undependable ☐
COMPATIBILITY Attitude Towards the Company Attitude Towards Supervision Co-operation with Fellow-Employees	Inspires others to work with and assist co-workers ☐	Quick to volunteer to work with and assist others ☐	Generally works well with and assists others ☐	Seldom works well with or assists others ☐	Does not work well with or assist others ☐

COMMENTS ____________________

FIG. 112—A typical merit rating chart used in industry.

more employees—the process of ranking becomes unwieldy. This unwieldiness is due partly to the mechanical difficulties

of ranking a large group of men and partly to the fact that when many men are involved there will frequently be several who, in the opinion of the supervisor, should be given equal ratings. To meet this situation, the forced distribution system provides that employees be rated on one or more scales, with approximate percentages of employees stipulated for each scale location. Thus, on a scale of "performance on present job," the following percentages might be used:

Lowest 10% *Next* 20% *Middle* 40% *Next* 20% *Highest* 10%

The use of a forced distribution of this type prevents some supervisors from using only the high part of the scale, others from using another part exclusively. The percentages should be considered as guideposts rather than as rigid rules.

The forced distribution system can be used when ranking a number of traits, if that is desired; in this case, the employees may be separately rated on each trait. However, the evidence from statistical studies such as those reported on pages 333 to 340 is rather strongly in favor of rating employees on only a few traits. On the basis of this evidence, a simple, two-item, merit-rating card has been developed. On this card, reproduced in Figure 113, each employee is rated on "job performance on present job" and "supervisory possibilities." In the training program preceding the use of these cards, the raters are acquainted with the percentage distribution that should be used as a guidepost in ranking. It will be noted that no verbal descriptions are printed under the five boxes on the card in Figure 113. Verbal descriptions are not used because, if they were used, the description under the "Low" box would necessarily have to be rather unfavorable, and many supervisors object to placing a man in the low category when it is qualified by an unfavorable comment.

A second rating on "supervisory possibilities" is also called for by the card shown in Figure 113. The use of forced distribution in rating this trait is not recommended, because in

some departments there may be no employees who are potential supervisors, while in other departments several may have potential supervisory ability.

In using a two-characteristic rating system of this type, the two ratings should not be added together to give a single composite rating. To do so obscures the significance of the individual ratings.

If the ratings are to be used for any purpose that requires discussing them with the men rated, it is usually desirable to

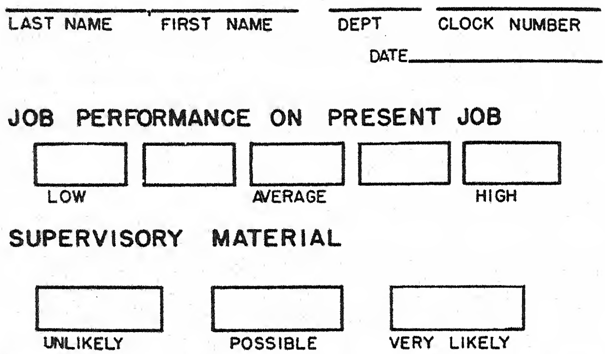


Fig. 113—Form used in merit rating employees by a forced distribution system.

have check lists of characteristics that justify the overall rating made. The list may consist of a set of traits such as those tabulated in Table 27, with corresponding boxes to be checked for those traits in which the employee is strong and for those in which he is weak. These lists may be printed on the back of the card illustrated in Figure 113.

Some Values of Merit Rating

Regardless of a number of criticisms, statistical and otherwise, that may be directed against merit rating—and that will be discussed on page 333—industrial relations managers are coming to recognize, more and more, certain values inherent in such a system. To obtain many of these values, however, it is necessary that the results of the merit rating be

made known, through the proper channel, to the employees who have been rated. Whether the "proper channel" should be the supervisor who made the rating, the superintendent of the department, a psychological counselor, or someone else, depends upon a decision as to who has the time and training to assume this responsibility. Usually it is advisable for *someone* to discuss the ratings with the men rated in order, as Davis[6] has pointed out, to avoid misinterpretation. Moore[7] has remarked that one of the most valuable uses of a rating program is to identify the weaknesses of employees that may be corrected. Knowles[8] also has emphasized that employees should be informed of their ratings and encouraged to overcome their shortcomings. Armstrong[9] has outlined several principles that may be advantageously followed in discussing ratings with employees, namely:

1. Criticism and reprimand should be based upon facts and should avoid personalities.
2. Public criticism is rarely justified.
3. Private reprimands should be preceded by a statement of some of the more desirable traits of the employee.

The experience of many industries has been that employees should be informed of their ratings by some production man—either the foreman himself or the departmental head. Production men, however, often require special training before they are able to do this smoothly and effectively. Many supervisors who are good production men are not very skillful in discussing a man's weak points face to face with him. To do so successfully calls for tact, objectivity, and a sincere interest in helping the man as well as the business. Many

---

[6] H. A. Davis, "Inefficient Efficiency Rating," *Personnel Journal*, XXII (1944), pp. 268–270.

[7] H. Moore, "Real Use for Rating Scales," *Personnel Journal*, XXI (1942), pp. 165–170.

[8] A. S. Knowles, "Merit Rating and Labor Management," *Personnel*, XVII (1940), pp. 29–42.

[9] T. O. Armstrong, "Talking your Ratings," *Personnel*, XX (1943), pp. 112–115.

industries have found it wise to devote an extended series of training conferences to teaching supervisors how to inform their men of weak spots without offending the men. When this result has been accomplished and when a supervisor is able to talk face to face with his men about their weak as well as their strong points, a long step has been taken toward solidarity in the working group and toward the upgrading of men who otherwise might, for lack of information, either remain exactly where they are or actually regress.

Prevents grievances

A merit-rating system stimulates supervisors to talk over with their men possible sources of grievance before serious problems have a chance to arise. It is not only sound psychology, but good common sense as well, to "nip in the bud" any situation that may cause friction between two people. Many such situations are certain to arise in a working relationship as close as the one between employee and supervisor. Frankness on the part of the supervisor in making legitimate suggestions is facilitated by a merit-rating system. The spirit of constructive criticism can begin with the construction of the rating scale. Slocombe[10] has suggested that foremen and men to be rated should work together in deciding upon the content of the merit-rating chart. Such co-operation makes everyone concerned with the merit rating become a partner in the enterprise and paves the way for an effective and serviceable use of the system.

Improves job performance

If an employee, either for lack of experience or lack of information, continually performs his job in the wrong way, the supervisor will be doing both the employee and the company, as well as himself, a favor by bringing this matter to the employee's attention in the proper manner. However,

---

[10] C. S. Slocombe, "Psychology of Co-operation," *Personnel Journal*, XVI (1938), pp. 325–332.

unless a supervisor is required periodically to evaluate all of his employees with regard to their ability in various directions, it is quite likely that many aspects of an employee's performance that could easily be improved, and that the employee himself would like to improve if he were made aware of his shortcomings, will go on from day to day without correction on his part. To inform an employee of his strong and weak points is not only a reasonable and fair thing to do but is also good business for any management that is sincerely interested in having its employees perform their jobs in the best possible way.

Increases analytical ability of supervisors

It is very easy for a supervisor to judge a man as good or bad, strong or weak, desirable or undesirable, without asking himself the question, "Why have I judged this man in this particular way?" But when a supervisor asks himself the question, "Why is a certain employee unsatisfactory?" and when he is assisted in asking this question by a merit-rating system that lists a number of characteristics in which the employee may be strong or weak, it is often possible to identify immediately a particular weakness that accounts for all, or nearly all, of the employee's undesirability upon a certain job. Perhaps the employee is on a job for which he does not have the necessary dexterity, strength, or adaptability. When one has determined *why* an employee is unsatisfactory on a certain job it is often possible to shift that individual to some other type of job for which he is more adapted and on which his performance will be more acceptable. Merit rating increases the supervisor's consciousness of the need for such analysis of his employees.

Assists management in promotion, demotion, and transfer problems

Another value of merit ratings is that they help to supply an adequate basis for promotion, demotion, and transfer.

In any large industry it is necessary to make changes of this type. Often such changes must be made quickly in order to take care of new business or to adapt the personnel of a plant to technological changes. A systematic and periodic record of an employee's rating while he has been on various jobs in the organization greatly simplifies the shifting made necessary by requirements of promotion, demotion, and transfer.

Union contracts often place great emphasis upon the principle of seniority as a basis for job changes of employees within the jurisdiction of the contract. Under these circumstances, management's judgment, based on merit ratings, often cannot be exercised. But some job changes, such as promotion to supervisory jobs, often need not be based on seniority, and in such cases as these, systematic merit ratings often provide a valuable tool in identifying the proper men for promotion.

Reveals areas where training is needed

A fifth value of merit rating that is becoming more and more apparent with the growth of industrial "in-service" training is the manner in which such ratings help to locate areas where training can be given to advantage. For example, it is not uncommon to find that in a certain department the average rating of all employees on such a factor as "knowledge of their job" is consistently lower than the ratings of employees in other departments on this characteristic. This situation can often be remedied by the preparation of a job-training manual and the offering of training on those aspects of the job that will upgrade these particular employees. Likewise, a department in which the men are consistently rated lower in "co-operation" than are the other employees in the plant may bring to the attention of management a department in which something is wrong in the relationship between workers and supervisor. To determine whether such a situation is due primarily to the workers or to the supervisor often requires careful study. The difficulty

may be due to the fact that the workers cannot get along with the supervisor, in which case perhaps the latter should be changed rather than the workers reprimanded. Whatever may be the cause or the eventual solution, the presence of such a situation is often brought to the attention of management only after a systematic plan of merit rating.

Some Dangers of Merit Rating

Although any merit-rating chart or system should be adapted to the needs of the organization in which it is to be used, certain principles, if followed, will definitely increase the value of the ratings and, if neglected, will cause, at the best, a marked reduction in the value of the ratings and, at the worst, serious trouble in the form of labor difficulties or industrial relations disputes. Several of these principles have been the subject of extensive study by psychologists.

The halo effect

More than twenty years ago Thorndike[11] pointed out on the basis of experimental evidence that a rater has a constant tendency to rate an individual either high or low in many traits because the rater knows (or thinks) the individual to be high or low in some specific or particular trait. Thorndike called this tendency the "halo" effect. Applied to the industrial situation, Thorndike's statement means that if the supervisor regards an employee as very satisfactory in terms of his general personality and co-operativeness, he is likely to rate the employee high also in such traits as productivity, ingenuity, inventiveness, adaptability, and perhaps many other traits.[12] In other words, it is difficult for any rater—particularly an untrained rater—to isolate and rate separately the various traits that an employee may possess. We have

[11] E. L. Thorndike, "A Constant Error in Psychological Ratings," *Journal of Applied Psychology*, IV (1920), pp. 25–29.

[12] A discussion of the halo effect in the industrial application of merit rating is given by R. S. Driver, "A Case History in Merit Rating," *Personnel*, XVI (1940), pp. 137–162.

pointed out that the use of a merit-rating chart is likely to increase the ability of a foreman to make an analytical judgment; but we know that even under the most favorable conditions the halo effect will be present to some extent and that its results will be most prominent where ratings have been made by those unfamiliar with its very existence.

The halo effect can be minimized in several ways. If a chart system is used, it is usually advisable to follow the suggestion of Stevens and Wonderlic[13] and have each supervisor rate all his men on one trait before going on to the second trait, on the second trait before going on to the third, and so on. Since this method causes the supervisor to think of all his men in connection with a given trait alone rather than to think of each man as a whole, the effect of this general change in point of view is a reduction of the halo effect. This effect can also be minimized by arranging the chart itself so that the desirable end of some traits is on the right-hand side of the scale whereas the desirable end of other traits is on the left-hand side of the scale. This procedure prevents a supervisor from checking down a column on the right-hand side for a generally desirable employee, or down the left-hand side for a generally undesirable employee.

The operation of the halo effect in an actual set of ratings is shown in Figure 114. This figure reveals graphically, for 18 randomly selected men from a large industrial organization, the relationship between ratings on overall job performance, accuracy, safety, and co-operation, on the one hand, and on the other, an overall merit rating which included the four traits mentioned above along with eight others that need not be identified. The 18 men are arranged in rank order from left to right according to their overall merit rating. These ratings are indicated by the heavy black line. The ratings of the 18 men on the other four traits are indicated by the

[13] S. N. Stevens and E. F. Wonderlic, "An Effective Revision of the Rating Technique," *Personnel Journal*, XIII (1934), pp. 125–134.

remaining lines. It will be seen from an inspection of Figure 114 that those individuals who tend to rate high in any given trait, say co-operation, are also rated about equally high on all of the other traits shown. Of course, it may theoretically be

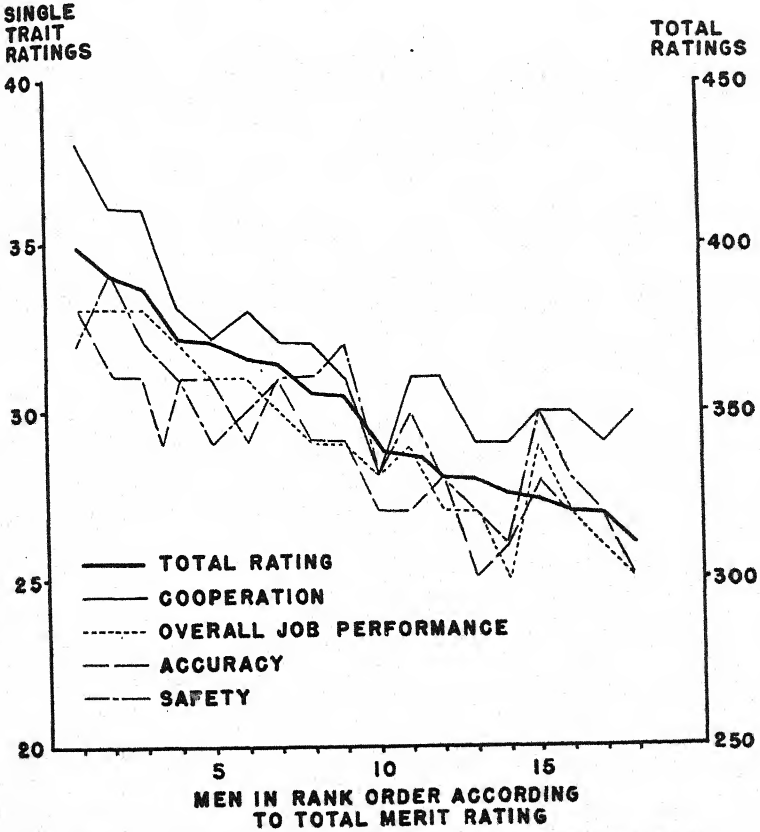


FIG. 114—Operation of the "halo effect" in industrial merit rating.

possible that the individual high in one trait is actually high in all of the other traits, but it seems much more likely that the halo effect, rather than any real relationship among the traits, is operating here.

Further evidence of a more extensive statistical sort of the

existence of the halo effect has been obtained from a factor analysis of approximately 1100 ratings selected at random from a plant in which about 9000 employees had been rated.[14] The chart used involved ratings on twelve separate traits. These traits and the intercorrelations between the ratings for each pair of traits are shown in Table 28. Such a correlational matrix containing correlations in general as high as those shown in Table 28 is sufficient in itself to indicate that a common factor is operating to influence most, if not all, of the ratings. The results of a factor analysis of these ratings quite definitely justify this judgment. Factor analysis is a

TABLE 28
INTERCORRELATIONS OF THE TRAITS IN A TWELVE-ITEM MERIT-RATING SCALE

Traits	1. Safety	2. Knowledge of Job	3. Versatility	4. Accuracy	5. Productivity	6. Overall Job Performance	7. Industriousness	8. Initiative	9. Judgment	10. Co-operation	11. Personality	12. Health
1. Safety												
2. Knowledge of Job	.61											
3. Versatility	.52	.81										
4. Accuracy	.63	.85	.80									
5. Productivity	.55	.79	.72	.81								
6. Overall Job Performance	.60	.82	.80	.67	.86							
7. Industriousness	.49	.78	.71	.80	.86	.85						
8. Initiative	.54	.78	.78	.78	.80	.83	.82					
9. Judgment	.62	.80	.82	.84	.81	.88	.84	.86				
10. Co-operation	.61	.67	.68	.74	.81	.80	.80	.72	.76			
11. Personality	.55	.67	.63	.70	.73	.74	.67	.72	.75	.80		
12. Health	.25	.52	.50	.84	.45	.60	.53	.77	.43	.52	.71	

[14] Edwin Ewart, S. E. Seashore, and Joseph Tiffin, "A Factor Analysis of an Industrial Merit Rating Scale," *Journal of Applied Psychology*, XXV (1941), pp. 481–486.

statistical technique that reduces a set of measurements (such as test results or merit ratings) to the minimum number of basic variables or factors that will account for the variations in the original data. The factor analysis of the merit ratings revealed three basic factors, factor loadings for which are given in Table 29. In interpreting the importance of the different merit-rating items as they enter into each of the three factors, we should bear in mind that the amount of

TABLE 29
FACTOR LOADINGS AFTER ROTATION

	Factor Loadings			h^2*
	I	*II*	*III*	
1. Safety	.633	.245	−.039	.462
2. Knowledge of job	.841	.337	−.002	.821
3. Versatility	.795	.325	.030	.739
4. Accuracy	.826	.448	.297	.971
5. Productivity	.913	.039	−.042	.837
6. Overall job performance	.961	−.080	−.064	.934
7. Industriousness	.906	.009	.007	.821
8. Initiative	.887	.094	.252	.859
9. Judgment	.897	.265	−.036	.876
10. Co-operation	.881	−.025	.042	.779
11. Personality	.815	.009	.259	.731
12. Health	.573	−.016	.836	1.027

* Communalities.

the factor loading indicates the extent to which that item is related to the factor in question. An inspection of the merit-rating items entering into Factor I reveals that overall job performance is loaded the most heavily; productivity, second; industriousness, third; judgment, fourth; and so on down the list. If we were to attempt to name this factor, probably the best name available would be "Ability to Do the Present Job." In other words, this factor seems to be a job-performance factor. Since the factor loadings on all of the traits entering into Factor I are high, compared with the loadings

on Factors II and III (except Health in Factor III), we infer that this factor is by far the most important of the three in determining the overall merit rating and that, since all of the heavily loaded traits entering into Factor I deal rather specifically with ability on the job, it would seem safe to identify this factor as a job-performance factor.

Factor II contains no elements so heavily loaded as are those entering into Factor I. However, those most heavily loaded in Factor II are accuracy, knowledge of the job, and versatility. Judgment and safety are the fourth and fifth, respectively, but they are not so highly related to Factor II as are the first three traits named. If we attempt to name Factor II, a name must be found which combines the five traits as they are all related to the unnamed factor. Since none of the loadings for this factor is high in comparison with those for Factor I, we must conclude that none of the twelve items of the rating scale "taps" this factor very well. We are faced then with the task of naming this factor although we know only the following facts: (1) accuracy correlates with it .45, (2) knowledge of the job correlates with it .34, (3) versatility correlates with it .33, (4) judgment correlates with it .27, (5) safety correlates with it .25, and (6) no one of these items from the rating chart is a satisfactory name for the factor because no one of them correlates highly with the factor.

From this evidence it would seem that Factor II may deal with the *quality* of performance on the job. Skill and accuracy are also suggested as possible names. Since knowledge of the job has the second highest factor loading, perhaps promotability or likelihood of being promoted would be an appropriate identification of this factor. Whatever we may choose to call it, certainly this second factor is not the same as Factor I, and it does not have as much influence on the total ratings, under the system which had been in use when these merit ratings were obtained, as does Factor I.

Factor III was found to be significantly loaded with only one element, namely, health. In other words, this factor

was related to the health ratings and to nothing else. It should be mentioned at this point that the reliability of the ratings of the various traits had previously been obtained and, of all the traits, the one rated with the least reliability was *health*. Indeed, the reliability of the health ratings was only .37, and yet here we are presented with a situation in which the factor analysis shows a trait whose internal reliability is only .37 to have a factor loading on the third factor of .84. This would mean that a trait that correlates with itself only to the extent of .37 nevertheless correlates with something else (Factor III) to the extent of .84. While it is possible for the correlation of a set of measurements (such as ratings or test scores) with an external measure to be higher than the reliability of the original measurements, it is difficult to explain so great a difference as the one found here. The answer to the paradoxical situation is that the unknown factor identified in Factor III can be nothing but the unreliable health ratings themselves. The analysis postulates a third factor to account for these health ratings for the same reason that any unreliable, inconsistent, and random measurement would necessarily come out in a factor analysis as an independent factor. Such a factor would not significantly correlate with any other set of measurements.

This evidence provides quite sufficient grounds for disregarding the health ratings as indicating anything except random and relatively worthless judgments of the moment. A further reason (if one were needed) for disregarding the health ratings would be the fact that an employee's health is, of all things, something that should be determined scientifically by a physician and not by the judgment of his foreman or supervisor.

The logical conclusion one would reach from the foregoing discussion is that the merit-rating chart in question should be reduced to only two traits and that these traits should be identified so as to cover as well as possible the two factors that the analysis revealed. This conclusion is the basic reason for

the development of the simplified forced distribution system described on page 325. Such a reduction in number of traits rated involves no sacrifice in either reliability or accuracy. However, it may well be that certain other advantages that accrue from merit rating—such as the increase in the analytical viewpoint of the foreman and the identification of the employee who needs help or training—can be better achieved by a more detailed type of merit-rating chart. The advantages of the analytical approach, however, can be obtained by the adoption of the check lists recommended on page 328 for use with the forced distribution system. The decision as to which system is to be used must be made by each industry in the light of the evidence and after a careful consideration of the needs of the particular plant.

It should be stated that the presence of only two factors in the ratings on twelve traits may be a reflection of inadequate training of the supervisors in the general principles of rating employees. A study by Driver[15] of a similar set of ratings obtained from supervisors who had been given seven hours of intensive training in rating methods revealed much lower intercorrelations than those shown in Table 28.

The general conclusion with regard to the halo effect is that unless supervisors have been trained very carefully they may allow one trait of an employee greatly to influence their ratings of that employee on many other traits.

The "weighting" of traits

After a decision has been reached in regard to the particular traits that are to comprise a merit-rating chart, many industries have felt it desirable to weight each of these traits according to what seems to be the relative importance of each for success of employees in that particular organization. Thus, one industry might weight initiative 20 points and safety 5 points, whereas another organization might reverse

[15] Unpublished Study by R. S. Driver, Atlantic Refining Company, Philadelphia.

this weighting of items. Table 27 shows that of the eighteen charts analyzed, four, or 22 per cent, incorporated some sort of differential weighting for the various items. The remaining charts use either equal weightings or no weightings at all. Either of these methods is presumably intended to give the same importance to the several traits included on the chart.

At least two difficulties arise in connection with weighting of items on a merit-rating chart. The first of these is that items are not necessarily weighted equally when all are given the same maximum numerical value, nor are they necessarily weighted in the way intended when a predetermined set of maximum values for each is used. In combining scores—whether they are merit-rating items, test scores, production records, or any other set of values—the scores weight themselves automatically in proportion to their respective variability. Expressed in statistical terms, the scores weight themselves in proportion to their respective standard deviations. Therefore, if the variability of all employees on one rating item, say health, is twice as large as the corresponding variability of all employees on some other item, say initiative, a direct combination of ratings for any employee on these two traits is actually weighting the health ratings twice as heavily as the initiative ratings. It might, of course, be the wish of management to weight those two items in this proportion, but it is unlikely that the chance weights that creep into a set of ratings as a function of their respective variabilities will weight the various traits in the manner desired by management.

The statistical reason for weights being determined by variability of the ratings is discussed in some detail in Appendix A, page 501. An example here may further clarify the principle.

Suppose that 1,000 men have been rated on two traits, namely, health and initiative. Each man has been rated on each trait on a 50-point scale. Suppose, for the present illustration, that all of the men have received ratings on

initiative of between 30 and 35 points. Suppose, further, that the health ratings vary from 25 to 45. If we now combine for each man his rating on initiative and his rating on health we will obtain a combination rating in which it has often been assumed (because each trait was originally rated on a 50-point scale) that the two ratings are weighted equally. Under these circumstances, however, the traits are not weighted equally at all. The health ratings, which vary over a range of 20 points—from 25 to 45—will have approximately four times as much effect on the total rating as the ratings

TABLE 30

STANDARD DEVIATIONS OF RATINGS OF EMPLOYEES IN TWELVE TRAITS[16]

Trait	*Standard Deviation*	*Relative Weight*
1. Safety	2.24	1.00
2. Knowledge of job	2.77	1.24
3. Versatility	2.88	1.29
4. Accuracy	2.69	1.20
5. Productivity	2.58	1.15
6. Overall job performance	2.63	1.18
7. Industriousness	2.96	1.32
8. Initiative	3.08	1.38
9. Judgment	2.68	1.20
10. Co-operation	2.72	1.22
11. Personality	2.51	1.12
12. Health	3.14	1.40

on initiative, which vary over a range of only 5 points—from 30 to 35. A method of combining the ratings so that they may be weighted equally is described in Appendix A. If this method seems rather complicated, it can only be said that the rating of human personality traits is complicated and is difficult enough to accomplish satisfactorily even when all statistical safeguards are followed.

The fact that unknown weights for the various items on a merit-rating chart not only may but *do* creep in, if not guarded against statistically, is proved by an analysis of the merit ratings of 1000 employees in a steel mill. The variability

[16] This table is from J. Tiffin and W. Musser, "Weighting Merit Rating Items," *Journal of Applied Psychology*, XXVI (1942), pp. 575–583.

of the ratings on each of twelve traits was computed and the standard deviations are shown in Table 30.

The employees had been rated on each of these traits on a 50-point scale, and it was assumed that this procedure resulted in total ratings that were influenced in an equal amount by each trait. Actually, the health ratings, which had the largest standard deviation, were exercising 40 per cent more effect on the total ratings than the safety ratings, which had the smallest standard deviation. Next in importance were the ratings on initiative, which were exercising 38 per cent more effect than the safety ratings. The relative weights actually exercised by each of the twelve factors are given in the final column of Table 30. It is doubtful whether the management of this company would have agreed upon this weighting of items if the matter had been discussed when the chart was constructed.

The simple adding of ratings for several traits not only fails to weight the traits equally (except occasionally and by chance) but also fails to give them any preassigned weights that might have been decided upon and crystallized in the form of a maximum value that each trait may receive. Suppose, for example, that management has decided that accuracy is twice as important as production and therefore has adopted a chart in which accuracy is rated on a 40-point scale and production on a 20-point scale. This arrangement will not necessarily result in accuracy being weighted twice as heavily as production, for *the relative weights of the traits are determined by the variability or spread of each and not by the maximum values assigned to each.* It would be quite possible in the situation described above for the production ratings to assume the heavier weights if the original ratings were directly added.

A second hazard related to the matter of weighting items on a merit-rating chart, even when proper steps have been taken to insure the functional operation of the weights

decided upon, deals with the differential weighting of traits for employees on different jobs.

Suppose that a system is set up to weight experience 20 points and education 5 points. Let us say that an older man who has been on his present job for many years is now to be compared with a research man who has recently been employed. The older employee, because of his extensive experience, will receive practically the maximum amount of rating points for experience, but, because of lack of education, will receive little on that trait. Let us say that he gets the full 20 points for experience but only 1 or 2 points for education, giving him a total merit rating of 21 or 22 points. The research worker, on the other hand, having only recently joined the organization, receives a minimum of points on the experience rating but the maximum of points on the education rating. Let us say that he receives 3 points for experience and the maximum, or 5 points, for education. His total merit rating would therefore be 8 points which, compared with the 22 points received by the older employee, would seem on the surface to indicate that the older worker is a considerably more valuable individual to the organization than the research worker. Now it is entirely possible that the older worker is of more value to the company than the research worker, but it is by no means certain that he is. The point being made here is that a merit-rating system that weights the items as we have mentioned above is almost certain to give the worker of longer experience a higher total rating than the younger man.

It may be suggested that this difficulty may be eliminated by weighting all items equally. However, this procedure assumes that for all jobs on which the merit-rating system is to be applied the various traits included in that merit-rating system are of equal importance. This plan also involves an error in many situations. Consider, for example, the case of rating two clerical workers, one a receptionist secretary and one a private secretary to the plant superintendent. In many

organizations the receptionist secretary, in order to do her job efficiently and satisfactorily, needs a maximum of appearance, poise, tact, and friendliness, and very little, if anything at all, of such traits as originality, versatility, and intelligence. The private secretary, on the other hand, needs all of these latter traits, although she perhaps does not need to be quite so personable on first appearance as does the receptionist secretary. Several traits might well operate to give a total overall rating to the receptionist secretary considerably lower than that of the private secretary; and yet this receptionist secretary might possess the one or two characteristics that such a job requires, so that not only is she highly satisfactory but actually she is able to do this job even better than the private secretary would be able to do it if she were transferred to it. In this connection Locke[17] has suggested that the merit rating system should be keyed to the jobs concerned.

The use of a single overall trait called job performance as a criterion, as recommended on page 327, largely eliminates this problem. A supervisor in rating an individual in terms of his or her ability to do his present job tends to take all of these factors into consideration in an unconscious general evaluation. In the situation referred to above, it is quite likely that the receptionist secretary would be rated in terms of her present job as highly as the private secretary is rated in terms of hers. But if the merit-rating chart is so set up that it is necessary to rate each individual on a number of specific points and then to determine the total merit rating from some combination of these several ratings, it is quite likely that an overall result not compatible with the facts will be obtained. Here, then, we are presented with a situation in which we must be careful about judging an employee's fitness on his present job by adding the results of several merit-rating items. From this point of view, the halo effect, as Bingham[18] has pointed

[17] N. Locke, "Employee Ratings," *Personnel Journal*, XXI (1943), pp. 282–288.

[18] W. V. Bingham, "Halo, Invalid and Valid," *Journal of Applied Psychology*, XXIII (1939), pp. 221–228.

out, may be the saving grace of a fractionated merit-rating system. In other words, when a supervisor begins to rate any given employee on several traits, he may, and probably does, first of all center his attention upon the particular trait or traits that are necessary for the employee on his particular job. After these traits have been rated, all of the remaining traits on the chart automatically fall in line in terms of where the key traits have been rated. But the old axiom that one cannot eat his cake and have it too prevails in the use of an itemized versus an overall merit-rating system. To the extent that the merit-rating system really does fractionate an employee's characteristics, it is unlikely to give an accurate indication of his ability on his present job; whereas to the extent that the system fails to fractionate the traits, owing to the halo effect, it is likely to work well in terms of giving an adequate overall indication. Some industries have solved this problem by using a dual system, that is, one in which the employee is first rated on overall ability on his present job before ratings of a fractionated sort are obtained. Such a dual system will accomplish both results with little more labor on the part of supervisors than is included in any of the systems now in use.

Pooling unreliable with reliable ratings

It has been shown by psychological research that the reliability of ratings is increased when it is possible to pool the ratings of several raters. Such pooling, however, assumes that the various raters are all competent to rate the employee in question. It is not necessary that the raters be equally competent or equally familiar with the employee, but it is necessary that they have enough knowledge of the employee so that their rating is not based largely on chance. Acting on this premise, many industries have installed a system that involves pooling of two or three or even more ratings before a final overall rating of an employee is obtained. However, the very organization of a modern industrial establishment is

such that the more ratings one attempts to obtain on a given employee, the further away from that employee in terms of contact with him one must go in order to obtain the additional raters. Unfortunately it is not necessary to go very far before one has brought in raters who are so far away from the actual observation of the employee on his job that their ratings of him are not only relatively worthless in themselves, but, what is worse, are also so inaccurate that when averaged with the ratings of the foreman in direct contact with the worker the result is to decrease whatever validity the foreman's ratings may possess. One rotten apple may spoil a bushel of good ones, and it is even more likely to spoil one or two good ones if those one or two happen to be in close contact with the spoiled one. Thus there is considerable danger in assuming that, because pooled ratings are known to have more reliability than individual ratings, pooling will automatically increase the validity of merit ratings. This result will not be accomplished if the pooling necessitates pooling unreliable with reliable ratings. An adequate merit-rating system should make provision for the rater to state how well he is acquainted with the employee and under what circumstances he has had an opportunity to judge him. This makes it possible to pool ratings when such pooling is likely to increase their value, and to avoid the pooling if by so doing the ratings would be made less valid.

Failure to determine the reliability of ratings

The general concept of reliability has been discussed in some detail on page 62 in connection with industrial-selection tests. It should be kept in mind that this concept is also directly applicable to merit rating. Only to the extent that repeated ratings will tend to give an individual the same rating, assuming that in the meantime he has had no opportunity to change, are we justified in allowing those ratings to influence our judgment of employees. The implication of this statement is that the more unreliable the ratings are, the

greater must be the change in an employee from one rating to another before we are justified in assuming that the change in rating actually indicates a corresponding shift in the employee's merit. One study by Reymert and Kohn[19] revealed the reliability for a nine-item scale to be .59 when two judges were used. This figure is rather typical of the reliability that may be expected of any merit-rating system based on two or three competent raters. When a ranking procedure rather than a rating chart is used in making the ratings, higher reliability (sometimes as high as .85 to .95) has been obtained.[20]

Data on the reliability of ratings on each item of a twelve-item rating chart, as well as the reliability of the total rating, are given in Table 31.

TABLE 31

RELIABILITY OF EACH ITEM OF A TWELVE-ITEM MERIT-RATING SCALE, AND TOTAL RATING BASED ON THE SUM OF ALL TWELVE ITEMS

Trait	*Reliability*
1. Safety	.35
2. Knowledge of job	.46
3. Versatility	.47
4. Accuracy	.45
5. Productivity	.46
6. Overall job performance	.46
7. Industriousness	.47
8. Initiative	.48
9. Judgment	.45
10. Co-operation	.37
11. Personality	.39
12. Health	.36
Total	.55

The reliability coefficients shown in Table 31 were obtained by correlating results from pairs of raters who had rated the same employee. The data are based on a total of 92 raters and 4,500 rated employees. While somewhat higher relia-

[19] M. L. Reymert and H. A. Kohn, "The Mooseheart Graphic Rating Scale for Housemothers and Housefathers," *Journal of Applied Psychology*, XXII (1938), pp. 288–294.

[20] Personal communication from Dr. H. C. Taylor of the Upjohn Foundation for Community Research.

bilities might be obtained in other industries by means of a different merit-rating chart, it does not seem reasonable to expect the reliability of any ratings based on only two or three raters to reach a much higher level than the figures here cited. It is unwise to place too much confidence in any measurement that has a reliability no higher than .55, the value obtained for the summation of items on the chart.

Another factor that often tends to make the ratings more consistent from one year to another than the facts of the situation justify is the supervisor's memory of his previous ratings of the employees. It is only reasonable to expect that when a supervisor rates a man twice, the second rating will be similar to the first if no marked new factors have arisen during the intervening period that might cause the supervisor to change his estimate of the man. The extent to which this memory element influences the consistency of ratings has been studied by determining the correlation between ratings on successive years for groups of employees who were rated by the same and different raters on the two occasions. The findings of this analysis are summarized in Table 32. An inspection of this table reveals a consistent decrease in the size of the correlation from the situation at the top of the list, where the same three raters were involved

TABLE 32

CORRELATIONS BETWEEN RATINGS ON SUCCESSIVE YEARS OF EMPLOYEES WITH SAME AND DIFFERENT RATERS

	Correlation
Same three raters both years	.65
One new rater second year	.59
Two new raters second year	.46
Three (all) new raters second year	.49

during the two successive years, to the bottom of the list, where two or three new raters were involved the second year. These correlations should not, of course, be interpreted as reliability coefficients because the ratings were separated by an interval of a year, for during a period of that length many employees might actually change in such a way as to justify

a change in the rating given them. However, the fact that a rather marked difference exists between the situation in which the same raters are involved and that in which one or more raters have been changed indicates that the constancy of the rater does affect the consistency of the rating.

All of these facts considered together indicate that merit ratings do not have a particularly high reliability. This fact does not mean that the ratings are of no value, but it does mean that one should be aware of their reliabilities and should not attempt to use merit ratings in a way that presupposes a higher reliability than they actually possess.

In proportion as ratings are unreliable it is not a valid procedure to consider a slight change in rating from one time to another as indicating a real change in the merit of the employee. By means of a simple statistical procedure, it is possible to obtain for merit ratings what is known as the probable error of measurement. Unless an employee's rating changes by at least four probable errors of measurement from one rating to another, it is unsafe to assume that any real change has occurred. In the case of the ratings that were used in determining the reliability coefficients summarized in Table 32, the probable error of the total rating was fifteen points: in other words, an employee must shift up or down by approximately sixty points before management is reasonably certain that an actual change in the employee's merit has occurred. It is therefore very important for management to know the reliability of the ratings that are used. Only through such knowledge is it possible to know whether the ratings indicate a real difference between two employees or a real change in an employee from one time to another.

Giving numerical values of ratings to employees

In one set of ratings studied, the numerical values, obtained by adding the points contributed by the twelve items comprising the scale, varied from 150 to 500. The reliability of this set of ratings as discussed in the preceding

section was such that two employees must differ by at least 60 points before it can safely be assumed that a real difference between the two employees exists. In the light of this situation it is unwise to give the employees exact numerical statements of their ratings. One learns early in life that the figure 400 is larger than the figure 399. Therefore, if two employees should receive ratings represented by these numbers, one is likely to feel elated and the other depressed, though there is no reason at all for assuming that such a slight difference represents any real difference between the two employees. It is much safer to divide the range of ratings into four or five categories and to tell each employee only in which category he is located rather than exactly where he stands in that category.

A second problem in giving out numerical ratings arises when an employee compares his rating with a rating given him at some previous time. Suppose, for example, an employee's rating in 1947 is 350. He is not satisfied and makes every effort during the following year to improve his performance on the job. Possibly he attends night school, reads material related to his job, and in other ways makes a sincere effort to upgrade himself. At the end of the year the ratings are repeated and he finds that this time his numerical rating is only 345. One familiar with the fact that ratings are none too reliable, even at the best, would certainly not consider such a slight decrease from one year to another as indicating a drop in the employee's merit. But the employee himself, if these numerical figures have been given out, is very likely to adopt the policy, "What's the use?" If he has done everything possible during the year to improve himself, and finds at the end of that time that he has decreased in value to the company, he is unlikely to make any serious attempt in the future to improve his ability. This difficulty can be largely eliminated by giving out only general classifications such as A, B, C, and D. A "B-grade" employee is less likely to experience a change of grade as a result of a few numerical points differ-

ence in his rating than if he is classified exactly according to his numerical standing.

Failure to consider departmental differences in rating

It often happens that the merit ratings turned in from different departments in a given plant differ markedly from one department to another. This discrepancy may be due

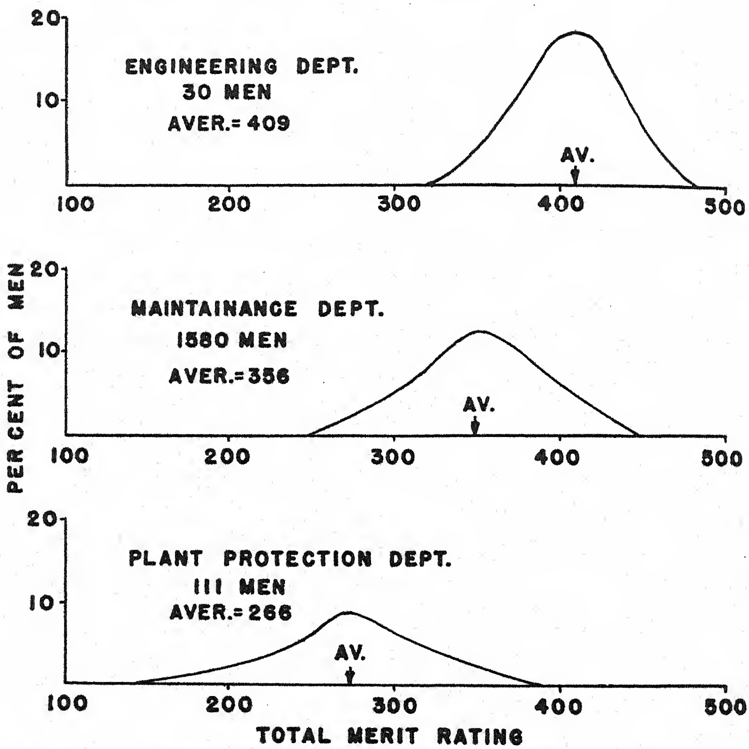


FIG. 115—Differences in merit ratings among departments in a steel mill.

in part to actual differences in the merit of employees in the various departments, but it may also be due in part to differences in standards or interpretation of the merit-rating scale among the departments. Whatever may be the cause, when such differences do occur it is usually desirable to evaluate a

given employee's rating in terms of the other ratings from the department in which he is working rather than in terms of the ratings obtained from the plant as a whole. This situation is graphically illustrated in Figure 115, which shows the distributions of total merit ratings obtained from three departments of a plant made up of 14 departments and employing approximately 10,000 men. The three departments selected for illustration are engineering, maintenance, and plant protection. These three have been selected to show the marked differences which may be found in ratings from one department to another. The difficulty of interpreting the significance of a merit rating without reference to the department from which it was obtained may be readily seen from inspecting Figure 115. For example, a merit-rating score of 350 would be a very low rating for a man in the engineering department. The same rating would be approximately average for a man in the maintenance department, and it would be very high for a man in the plant-protection department. Since one major function of a merit rating is to indicate how well—in relation to other employees—an employee is doing his present job, a fair and reasonable basis for comparison of ratings of different men must be employed. When ratings differ markedly from one department to another, evaluation of any rating should be in terms of the department from which it was obtained. The difficulty inherent in this problem can be largely eliminated by using separate norms for different departments.

Failure to consider job differences

Another source of possible difficulty closely related to the matter of departmental differences is the variation in rating often found from one job to another. When employees on any given job are consistently given higher merit ratings than are employees on other jobs, such job differences should be considered in evaluating the rating of any given employee. Figure 116 illustrates this situation for 51 jobs in a Sheet and

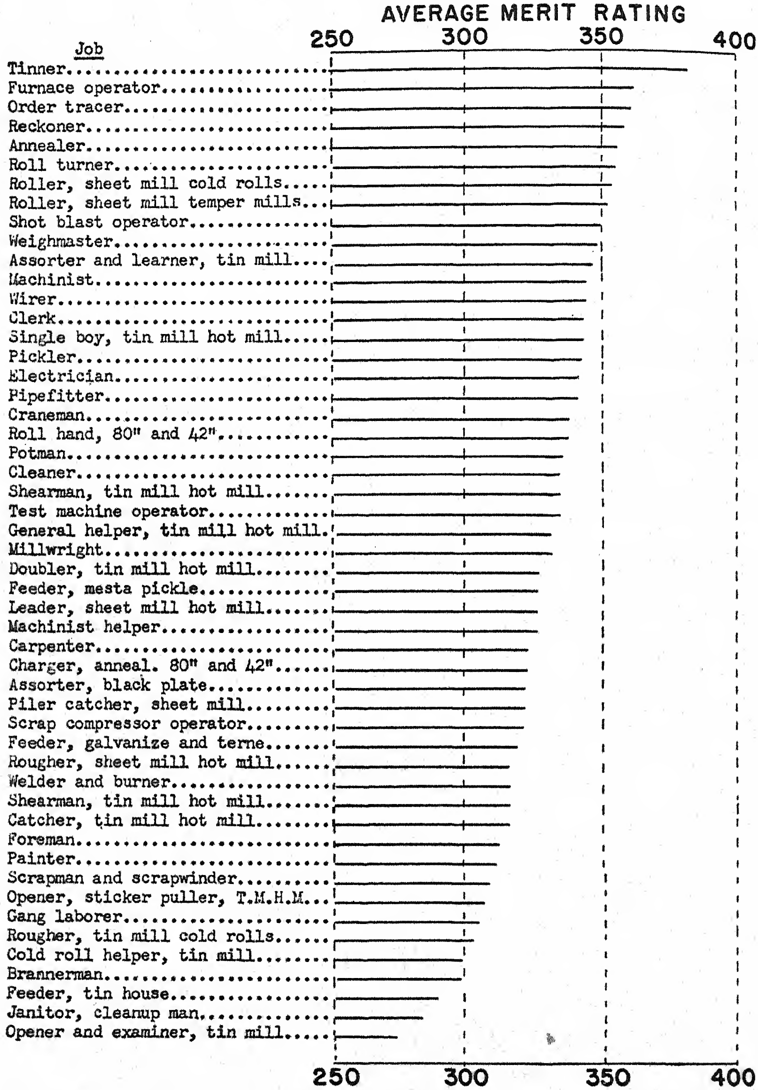


FIG. 116—Differences in average merit rating of employees on 51 jobs in a steel mill.

Tin Mill. The 51 jobs studied are arranged in order from the one receiving the highest average rating (tinner) at the top to the one receiving the lowest average rating (opener and examiner) at the bottom. It will be noted that there is a variation from 280 to 385, or 105 points, in average merit rating from the lowest to the highest average rating. From these differences it is clear that a rating of 300 is very high for an employee who is on one of the jobs located near the bottom of the list, but that 300 is very low for an employee on one of the jobs near the top of the list. The implication of this fact is that the merit rating of an employee should be evaluated in relation to the ratings of other employees on the same job or on jobs that are given approximately the same average merit rating.

Failure to consider the age of an employee

The age of an employee is another factor that is often related to the rating he receives. In one set of approximately

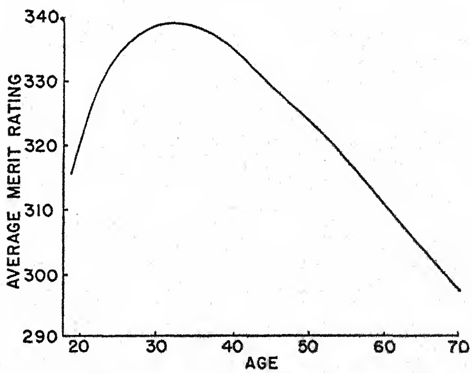


FIG. 117—Relation between age and average merit rating for 9,000 steel workers.

9,000 merit ratings obtained from a single industrial plant, the relationship between total rating and age is graphed (Figure 117). This figure shows that an employee between the ages of 30 and 35, other things being equal, is likely to get a higher

merit rating than an employee who is either older or younger. Figure 117 also shows that after this high point is reached a progressive decrease in merit rating takes place as age increases. It is clear from this chart that an employee with a merit rating of 330 would be definitely below average if he is in the age range from 30 to 35, but that he would be well above average if he is in the age range from 55 to 60.

Failure to consider other factors that may affect merit ratings

The preceding factors that should be considered in evaluating the merit rating of employees have been mentioned as

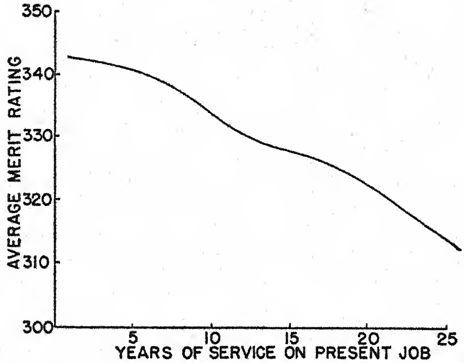


FIG. 118—Relation between years of service on present job and average merit rating for 9,000 steel workers.

illustrative of the kind of factors that have been found in a number of investigations to be related to merit ratings. Still other factors exist that may affect merit ratings in any given plant. Figure 118, for example, illustrates the relationship found in one plant between merit ratings and length of service on the present job. In definite contradiction to the opinion of management before these results were obtained, this chart shows a progressive lowering of merit ratings as the length of service on the job continues. In other words, the longer a given employee remains on a certain job, the lower the merit rating he is likely to receive from his super-

visor. Probably the explanation for this relationship is that only those employees who are ineligible for promotion or transfer to another or more important job are likely to remain on their present jobs for a long period of time. Or perhaps, in this particular plant, the employees with longer service on their present jobs are in general older employees and may be unable to do the work as effectively as younger men. Whatever may be the cause of the relationship, it is clear that it exists and that it should, therefore, be considered in evaluating a given employee's rating. A rating, such as 335, that might be well below average for an employee of

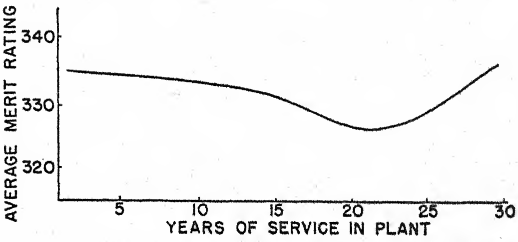


FIG. 119—Relation between years of plant service and average merit rating for 9,000 steel workers.

short service on the job, would be well above average for an employee who has been on the job 15 or 20 years.

Figure 119 graphs the relationship between merit rating and total service in the plant. It is clear that the drop-off or decrease in merit rating with total plant service is much less marked in the case of this relationship than in the case of the relationship with length of service on the present job. But here also a slight decrease in total rating seems to occur as the total plant service increases.

It is not meant to imply from the results summarized in the preceding sections that the relationships found in the studies reported exist in exactly this same form in all industries. Indeed, there is every reason to expect that these relationships are not universal and that opposite trends may occur in many plants. The point to be emphasized is not the

universality of the trends discussed but the fact that no company can properly interpret the results of the merit ratings in its own plant or plants without definite knowledge of what trends and extraneous factors influence these ratings in that particular organization. The only way an organization may determine this is to make a set of analyses similar to those reported in the preceding sections. Such a procedure may be objected to on the grounds that it would involve more labor, and perhaps expense, than merit ratings are worth. It is entirely possible that this is true. But, if one is to obtain merit ratings that are to be used in evaluating employees, it is not fair either to management or to the employees rated to use ratings when it is not known definitely what factors have influenced them. If it is not possible to determine what these several factors are in any given organization, it might be a wiser policy to discard the ratings altogether than to attempt to use them without this information.

Suggestions for Improving a Merit Rating Program

Many of the facts discussed in the preceding pages of this chapter may be put to practical use by deriving from them a number of suggestions for installing and operating a merit-rating system. Some of the more important suggestions are summarized below:

1. Train the Raters.[21] Rating people is a skill, and one that is not easy to learn. Good production men and even good supervisors are not necessarily good raters *unless they have been taught how to rate.* Even at best, rating is subjective and personal. Prejudice and bias can never be completely removed, and in the absence of training they may completely distort many merit ratings. Companies that have had all raters attend a systematic training program of from six to eight hours before making any ratings have found, without exception, that the time and money spent on this training was a good investment.

---

[21] For a more extensive discussion of the importance of training the raters, see R. S. Driver, "Training as a Means of Improving Employee Performance Rating," *Personnel*, XVIII (1942), pp. 364–370.

Subjects covered in the training of raters include:

a) The "halo effect"—what it is, how it affects ratings, and how to avoid or reduce it.

b) The importance of basing ratings, whenever possible, on facts rather than on opinion. In rating a man's production, specific instances when he failed to meet expected production, for example, are always more convincing than the simple statement, "production below par." Very often, specific instances can also be cited in support of ratings on co-operation, judgment, quality, and so on.

c) The necessity for the rater to make up his own mind and to refuse to be influenced by others. When a man has been transferred into a new department, it is not uncommon for the reaction of his former supervisor to his work to accompany him, and so, perhaps, to prevent him from getting off to a fair, fresh start. If his new supervisor rates him on the basis of his former supervisor's judgment, it is possible or even probable that the rating will not adequately reflect his performance on his new job.

d) The importance of using the whole spread on each item or trait. If a forced distribution system is used, the meaning of the guidepost percentages should, of course, be thoroughly covered in the training. But even if a chart system is used, it is important for the raters to realize that the whole spread of the traits should be used if the system is to operate effectively.

2. In addition to training as such, another procedure that increases the accuracy of ratings is to have them made in conference or under supervision. Adoption of this system does not mean that supervisors confer with each other about their men while the ratings are being made, but only that the ratings are made while the raters are gathered for this purpose under the guidance of someone thoroughly familiar with the system. By the use of this procedure, several difficulties will be avoided. Careful thought will be given to the problem, and the hasty checking that often takes place if the supervisor is expected to "find the time" to make the ratings will be eliminated.

3. After the ratings have been made, be extremely cautious in comparing the ratings of men in different departments or on different jobs. The need for this caution is brought out on page 353. Unless a correction has been made for a man's job, it is unfair to compare his rating with ratings of men on different jobs.

4. Interpret the ratings in terms of actual job demands. A man

should not be penalized because a rating shows him to be weak in some ability that his job does not require.

5. Avoid the use of numerical values, which give to ratings an appearance of greater accuracy than they in reality possess. The reason for this caution is discussed on page 350.

6. It is ordinarily wise to omit from the chart factors concerning which objective information can be obtained without using ratings. If production records are available, it is better to use these in evaluating the production of an employee than to rely on a supervisor's judgment. The same thing is true of quality of work, if an actual record of rejects or waste material is available, and of health. A supervisor can judge whether a man *looks* healthy or *acts* healthy, but it is the job of a physician to determine whether he really *is* healthy.

7. Avoid pooling poor ratings with good ones. The use of *one* satisfactory rating is better than the pooling of many, if some of those pooled are likely to be in error. This subject is discussed on page 346.

8. A final suggestion relative to the most effective use of merit ratings based on the chart system, or indeed, of merit ratings based on any system, is to make use of all possible supplementary information to be obtained from other sources. A very great deal of information about the performance of employees often remains unassembled and unused in many plants. Information on accidents and hospital visits, absenteeism, training time and costs, production, quality, and so forth, frequently is available, and can be very helpful in interpreting or supplementing merit ratings. Job-information tests are also being found useful by a number of plants as one of the tools to be used in upgrading and promoting. In some companies the union has endorsed the use of tests for this purpose when merit ratings alone were not accepted. In a grievance hearing, an unsupported merit rating presented to an arbiter as reason for an alleged discrimination against an employee with seniority is usually very difficult to justify. But a merit rating supported by one or more test scores on job-information tests or by records covering production, quality, and so forth, often forms a persuasive and conclusive case.

Merit ratings serve their purpose most effectively when used in conjunction with all available additional information about the employee's performance on the job. Merit ratings used in conjunction with such additional information and in the light of possible sources of difficulty discussed in the preceding pages furnish a valuable aid in the administration of an industrial relations program.

11

Wages and Job Evaluation

It is unnecessary to point out to the readers of this book that wages constitute one of the most persistent problems in the whole field of labor-management relations. Since unions first expressed their demands, an insistence on a higher general wage level as well as on certain other wage adjustments has characterized a majority of labor-management disputes. The importance of monetary demands by organized labor has not lessened in recent years, even though more non-monetary demands are being made now than in the past. An analysis of 2,055 union demands made in a sample of 821 dispute cases coming before the National War Labor Board during the years 1942–44 revealed that 910, or 44 per cent, of the demands involved monetary issues. A partial summary of this work has been published by Tiffin and Lawshe[1] and is reproduced in Table 33. The results tabulated show something of the variety of ways in which monetary demands may be made. The results also show that the percentages of "attained" and "partially attained" demands (even during the operation of a nation-wide wage "freeze") are sufficiently high to indicate that labor's arguments were considered rather persuasive by the members of the War Labor Board.

These facts suggest that we should carefully scrutinize the field of psychology (as well as such fields as economics and sociology) in the hope of finding a body of fact, or perhaps a *technique* or *method*, that will help us to arrive at a fair solution

[1] Joseph Tiffin and C. H. Lawshe, Jr., "War Labor Board Decision Trends," *Personnel*, XXII (1945), pp. 78–83.

of labor-management monetary disputes. Unfortunately, there does not seem to be any single *method*, acceptable to both labor and management, that may be applied to give a clear-cut solution to all wage disputes. The problem is not like one in arithmetic or mathematics, where, in seeking to find the square root of 63, anyone who knows the *method* will arrive at 7.936 as the answer, because there *is* no method of

TABLE 33

National War Labor Board Disposition of 910 Monetary Demands Made by Unions in 821 Dispute Cases Coming Before the Board

Type of Demand	Per Cent Attained	Per Cent Partially Attained	Per Cent Attained or Partially Attained	Per Cent Refused
Wages, minimum rates	63	11	74	22
Wages, hiring rates	50	17	67	25
Wages, substandard rates	47	29	76	24
Bonus payments	46	18	64	36
Severance pay	44	13	57	31
Wage adjustments	44	28	72	27
Wages, inequalities	43	26	69	28
Wages, going wage rates	42	22	64	36
Wages, cost-of-living adjustment	38	29	67	30
Wages, incentive wage system	37	23	60	37
Wages, trainee rates	36	28	64	36
Premium wage rates	34	27	61	38
Wage differentials	31	33	64	36
Wages, job rates	23	15	38	54
Total	41	26	67	31

solution that is acceptable to both labor and management. But there are certain factors that have traditionally affected wage rates, and an understanding of some of the most influential among them will help not only in developing a harmonious labor-management relationship but also in showing to both parties the vital importance of accepting some *systematic* method in the setting of wage rates.

Factors affecting wages can be divided into two broad (and sometimes overlapping) classes: namely, those affecting

general (i.e., community and industry) wage levels and those affecting wage rates for different jobs within a plant.

Factors Affecting General (Community and Industry) Wage Levels

Supply and demand

When no governmental or social restrictions on wages are in effect, the law of supply and demand has traditionally exercised a marked effect on wage levels. When labor is abundant, wages decrease; when labor is scarce, wages rise. Many managements which formerly allowed the operation of this principle to affect their wage structure have recently become more and more reluctant to follow the dictates of this "law." By paying an extremely high wage at times when labor is scarce, management has found that it has had trouble in meeting cost standards, and has been faced with a real difficulty in explaining the reason for wage reductions when the inevitable cut-back becomes necessary. And, conversely, it has found that the hidden costs in the form of employee dissatisfaction and pent-up resentment that result from the payment of a low wage when labor is abundant often overbalance in the long run the savings in wages. In general, management as well as labor is becoming less and less dependent on the so-called "law" of supply and demand as a basic factor in determining wage rates.

Governmental wage controls

Governmental wage controls fall into two major types, minimum rate and maximum rate restriction. The first type is illustrated by the Fair Labor Standards Act of 1938. This law provides that certain minimum wages (the minimum varying with the industry, locality, and date) must be paid by all companies engaged in interstate commerce or in the production of goods for interstate commerce. The setting of maximum wage rates by governmental order is illustrated by the directives of the Wage Stabilization Section of the

National War Labor Board during World War II. Maximum rates were set to prevent the monetary inflation that would inevitably have resulted if wages had been unlimited while most of the productive facilities of the country were engaged in production for war rather than for civilian consumption.

Civic relations of the company

Some companies have traditionally paid a higher wage than the prevailing rates in their locality in order to foster the continued good will of workers in the community. A company cannot, of course, follow this policy unless it is in a competitive position in the manufacture of its product that will permit such expenditures for wages. Companies that have been able to follow this policy have frequently built for themselves a *status* in the community that has minimized certain problems in employment experienced by other companies in the locality.

Cost of living

This is a factor that has been considered systematically by some companies in determining general wage levels. Yoder[2] discusses the use of cost-of-living indices for this purpose, and mentions several companies that employ this method of wage adjustment. He also analyzes a number of cost-of-living indices and discusses the advantages and limitations of each as a basis for general wage adjustments.

Regional and industrial differentials in workers' earnings

Closely related to figures on cost of living are variations in the general level of workers' earnings in different communities or localities. Workers living in cities where rents are relatively high, the cost of transportation to and from work is great, and other incidental expenses are proportionately large, are traditionally paid, by most companies, a

---

[2] Dale Yoder, *Personnel Management and Labor Relations* (Prentice-Hall, Inc., 1943), pp. 420–426.

higher general wage than workers with the same skills employed in smaller and more rural areas. This tendency has, at times, caused rural employees to feel that they are underpaid. However, in terms of "real wages," i.e., what the money will buy, it is sometimes true that urban workers are paid less in spite of their seemingly greater wages.

Strength of organized labor

When labor is organized and is directed by strong leadership, increases of general wage levels are sometimes obtained as a result of union-management negotiations. This statement does not mean that the strike, as such, is necessarily a device that should be thought of as the sole factor in determining wage levels. Any or all of the factors mentioned previously may operate as influencing factors, whether with or without the presence of organized labor. Nor does it mean that only through the strike is a responsible union leadership likely to exert any influence upon the wage policies of a company. The mere process of modern collective bargaining, as described by Hill and Hook,[3] sometimes results in a degree of understanding by management of labor's point of view that brings about certain wage adjustments, just as it sometimes results in a degree of understanding by labor of the problems of management that causes a withdrawal of certain wage demands.

Costs of production

Whatever may be a company's point of view toward any or all of the factors, mentioned above, that may have an effect upon the general wage level, the cost of production in relation to selling price must, of necessity, always set a ceiling on the wages that can be paid by any company that intends to remain in business. No company can follow for long a wage policy that requires a labor expenditure of $1.10 on a

---

[3] Lee H. Hill and C. R. Hook, Jr., *Management at the Bargaining Table* (McGraw-Hill Book Company, 1945).

commodity that must be sold (either because of competition or because of price ceilings) for $1.00. Obvious as this limiting factor may appear, it has not always been kept properly in mind by certain groups that have attempted to set wages in terms of one or more of the factors previously discussed without considering the selling price of the product.

The factors mentioned above have been discussed to bring out the fact that many elements, several of which are not an immediate concern of the psychologist as such, exercise an important influence in determining general wage levels. In any specific plant situation, several of these factors usually operate simultaneously, and are of varying relative importance depending upon the conditions operating at that time.

Factors Affecting Wage Rates Within a Plant

In addition to the factors that influence wages in general, certain additional factors operate predominantly to affect the relative wages of men or women on specific jobs within the plant.

Incentive versus hourly paid jobs

The earnings of employees working under an incentive wage plan (which provides that they be paid in proportion to their production) are often different from those of employees on the same or equivalent jobs who are paid by the hour. Indeed, incentive plans of wage payment are frequently installed for the purpose of increasing production, an increase which results in concomitant increases in the earnings of the employees. The installation of a suitable incentive plan for any given job or group of jobs is a matter of primary concern to the industrial engineer rather than to the industrial psychologist. Many variations of the basic incentive principle are found in the incentive plans of wage payment used by modern industry. Yoder[4] has summarized and described a

[4] Yoder, *op. cit.*, pp. 390 ff.

number of the more widely used incentive systems. A discussion of these plans would go beyond the scope of this book.

Non-financial incentives

Certain jobs are considered by many employees to be desirable for social or similar reasons, rather than for the wage paid to employees on these jobs. When this situation exists it is not uncommon to find the rates for such jobs set somewhat lower than are those for other jobs which require approximately the same level of ability, experience, or skill. Employees are often willing, or even anxious, to be assigned to such jobs because of the recognition or status involved, even though the actual monetary return is somewhat less than that from other jobs. Under the operation of any standard job evaluation system, such as will be described later, the importance of this factor in determining rates is usually minimized or even eliminated, but in the absence of a job evaluation system it sometimes exercises a marked effect upon certain rates within a plant.

Discrimination against women and other groups

Until quite recently certain plants followed a wage policy which permitted different rates to be paid to women on the same, or approximately the same, job as men. This practice was usually justified by the belief that women, because of their limited strength and physique, were unable or should not be asked to do all of the things that men on the same jobs were accustomed to doing. With the passage of recent social legislation and with the increased use of standard job evaluation systems, however, it has become necessary that a real difference in job requirements be demonstrated before a different rate can be set for men and women on the same job. Many inequalities formerly justified on this basis have therefore been eliminated by modern industry, and different

rates for men and women for the same type of work are less frequently encountered. In general, wage differences because of sex or other discriminative factors, though often encountered a decade ago, are becoming of less and less importance. This circumstance is due partly to legal and social changes, partly to the increased social consciousness of modern American management, and partly to a growing recognition of the fact that women and other groups formerly discriminated against can do many jobs in a highly satisfactory manner.

Company policy

A company going through the period of development intervening between a small organization and a large one usually finds itself compelled to place an ever increasing number of workers on a wide variety of newly created jobs for which specific and equitable rates have not, and could not have been, set in advance. In this stage of a company's development, there is seldom a single man or a department charged with the responsibility of developing and maintaining an all-inclusive wage structure that covers various employees on various jobs and in various departments. Many different supervisors in different departments hire men or women for various jobs, each one establishing within his own department rates that he considers fair and reasonable. Almost invariably the result of such a procedure is a set of rates for different jobs that are not properly and fairly related to one another. As the company grows and as these rates become the traditional rates by being paid over a period of time, it is not uncommon to find attempts made to justify the rate schedule on the grounds that the rates are in conformance with "company policy." This is only one of many illogical (and sometimes indefensible) practices that arise during the period of a company's growth and that must be remedied when that company reaches a size that requires integration and co-ordination of activities in various departments. A systematic job evaluation program, of the type that will be described later

in this chapter, results in the most reasonable and adequate solution of this problem.

Supply and demand

In addition to the effect that labor supply and demand have on the general wage level within a plant, these factors, at times, have influenced the rates paid for specific jobs. For example, if a number of welders are needed and there are few men available who are qualified for this work, the welder rate has sometimes been raised in an attempt to get more men for the job.

This method of "solving" a labor shortage in any particular classification is really no solution at all. It does not create more qualified employees. At best, it can do no more than "drain" men from other plants in the locality, and these other plants, by adopting the same procedure, can rapidly "drain" them back. In a very short time such an interplay will raise the rate for the job to an unreasonable level—unreasonable in terms of labor cost and also in comparison with other jobs requiring equivalent experience or skill. This latter point is perhaps the most persuasive reason for not allowing supply and demand factors to affect rates for specific jobs. Once a systematic and equitable rate structure has been installed, rates within established brackets should never be changed without a formal re-evaluation of the job. When fewer men for a given classification are available than are needed, it is advisable, both from the point of view of economy as well as from that of sound industrial relations practice, to institute a training or upgrading program to qualify men to fill the vacancies.

Collective bargaining agreement

Union contracts sometimes cover specific rates or rate ranges for specific jobs. This provision may result from the activities of a joint labor-management committee on job evaluation, or from a union demand for certain rates for

certain jobs when there is no formal job evaluation program in operation or when a program is in operation but the union has not been asked to participate in it.

A company can usually eliminate unreasonable union demands with respect to certain jobs by enlisting active union participation in the job evaluation program.

Job evaluation

A systematic evaluation of all jobs within a company is rapidly becoming the most widely accepted method of setting satisfactory and equitable rates or rate ranges for various jobs. Job evaluation is a subject to which several methods and techniques of psychology have been extensively applied. This topic will therefore be discussed in some detail in the section to follow.

Job Evaluation

One type of job analysis, as defined briefly in Chapter 2, is designed to form the basis for setting equitable rates for various jobs. This type of job analysis is referred to as job evaluation. Job evaluation is the rating of jobs according to a specific planned procedure in order to determine the relative worth of each job.

There are many reasons for the use of a systematic plan in setting wage rates. Without such a plan, rates will be affected by many of the factors previously discussed in ways unrelated to one another. Under such circumstances, the rate structure will contain many inconsistencies caused by rates that are not "in line." With a job evaluation plan in operation, inconsistency in rates is minimized and the entire wage structure becomes unified.

The installation of a job evaluation program involves several steps. Some of these steps are required by the mechanics of any system of job evaluation. Other steps, affecting primarily the industrial relations aspects of the installation, have been found expedient by various companies.

General Considerations in Installing and Operating a Job Evaluation Program

In installing and operating a systematic job evaluation program, several aspects of the project should be carefully considered in advance. Proper steps should be taken to anticipate difficulties, or at least to minimize them if they arise.

The job evaluation committee

Job evaluation requires not only technical competence in the field (including full knowledge of the system to be used), but also the active and enthusiastic co-operation of men representing several departments of the company. The departments most closely concerned are those concerned with industrial relations, personnel and employment, industrial engineering, and production. The job evaluation committee charged with the responsibility of installing the system should therefore be made up of men representing the interests of each of these groups. As to the advisability of having formal representation of organized labor on the job evaluation committee, there are different feelings by different managements. There is, however, rather general agreement that union representation on the job evaluation committee is advantageous so long as the recommendations of the committee are advisory and the committee is not given final authority on wage matters.[5] The basic reason for providing for union representation is that union members will very naturally want to know "what is going on." Unless some provision is made for their receiving direct, accurate, and comprehensive information about the project, they are apt to arrive at an entirely erroneous conception of its purpose. They may conclude, for example, that the jobs are being studied for the purpose of "cutting wages." It is obvious that the program will not function properly if initiated under such circumstances.

[5] Hill and Hook, *op. cit.*

Provision for evaluation of new jobs

No matter how thorough and satisfactory a job evaluation program is installed at the outset, provision must be made for continual, careful study and evaluation of new jobs as they are created. Jobs are not static because production processes are not static. The job evaluation committee should be a standing committee, and should meet at regular intervals in order to keep the evaluations current with the creation of new jobs and elimination of old ones as production processes change.

Getting jobs in line after the system is installed

A problem that is invariably encountered when a job evaluation system is installed for the first time is that some jobs will be found "out of line." These jobs will be of two types, those which have previously had too high a wage rate and those which previously have been set too low.

No problem is encountered in connection with the latter jobs, because the rate increase indicated by the evaluation will of course be accepted by the employees on these jobs. It is just as important, however, to fit those jobs which previously have been overpaid into the basic structure. It is obviously unwise, for many reasons, to adopt the policy of cutting these rates abruptly, and such a policy should never be followed (and almost never is followed) by a plant management. There are two recommended procedures to adopt to get overpaid jobs back "in line." First, every effort should be made to upgrade employees on such jobs to other jobs which permit them to maintain their present rate but which oblige them actually, to earn it. Such upgrading may at times require special training. When this is the case, it is advisable for the company to institute the training and make the upgrading as rapidly as possible.

A second step is to hire any new workers needed for these jobs at the rate determined by the job evaluation. If this second step is taken without the first, however, different

employees on the same job will be paid at different rates. Such a situation should be eliminated as rapidly as possible, and should not be allowed to exist at all except for a very short time while the transition is being made.

Obtaining the co-operation of supervision

In the past, the foreman has had many of the duties and responsibilities that have now been taken over by other representatives of management. Among these functions is the setting of wage rates. As a result, and as in the case of other changes in function and duties, many supervisors do not always recognize the necessity of a unified wage structure because the very nature of their respective jobs makes it difficult for them to see the problems of the company as a whole rather than just the problems of their own departments. It is therefore wise for management to devote as many supervisory sessions or conferences as necessary to discussions of the job evaluation program, to a full explanation of how the system works, and to the relation of the plant-wide program to the needs of the various departments.

Job descriptions for job evaluation

The first task of the job evaluation committee is to obtain adequate job descriptions. The importance of the latter as a basis for personnel specifications has been emphasized in Chapter 2. Adequate job descriptions are also of vital importance for a system of job evaluation. Jobs cannot be evaluated properly unless their nature is fully known. Just as job titles by themselves are of little or no value in writing personnel specifications, so also job titles must be accompanied by full job descriptions when the jobs in question are to be evaluated. Moreover, a given set of job descriptions is not always suitable for the various types of job analysis discussed in Chapter 2 unless care has been taken to include all information required for the different uses to be made of the descriptions. Therefore, when the job descriptions are prepared, it

is desirable to know (*a*) what they are to be used for, and (*b*) what information about the jobs is needed for all of the proposed uses. When this information is secured *before* the job descriptions are prepared, the time and cost of doing this work over at a later time are usually saved.

Job Evaluation Systems

There are several kinds of job evaluation systems in successful operation. These systems differ from one another in certain respects and it is always advisable, if at all possible, to decide on the specific system that is to be used *before* the job descriptions are prepared. By deciding on the system in advance, particular care can be exercised to include in the job descriptions whatever specific information may be required by the job evaluation system to be used.

Job evaluation systems have been classified in several different ways. Lytle[6] points out that they may be classified according to the *method* and also according to the *technique of measurement* used. This dual division of systems is perhaps more elaborate than is needed to summarize some of the basic plans now in use. We will therefore discuss certain plans which fall under four general groupings, as these cover the more widely used plans now in operation.

1. Scaling of jobs by comparing entire jobs

The simplest application of this method is through the principle of ranking. Under this plan, all jobs are ranked and placed in a continuum from highest to lowest. The ranking should be done by a person or committee familiar with all of the jobs. In a large plant, where several hundred jobs must be evaluated, this necessary familiarity is seldom possessed by any single person, or even by a group of persons. This fact is one of the most serious obstacles encountered in the use of the ranking system. It can be minimized to some

[6] C. W. Lytle, *Job Evaluation Methods* (The Ronald Press Company, 1946).

extent by having each rater rank only those jobs with which he is thoroughly familiar. Then, by a comparison of rankings given jobs common to several raters' lists, it is usually possible to establish key points in devising a scale that will include all jobs.

The reliability of rankings can be increased by using various psychological techniques, the value of which has been proved. Among these are the use of the method of paired comparisons (in which the rater compares each job with every other job) and the pooling of rankings of several raters to obtain average rankings.

2. Scaling of jobs by comparing components of each job

This method is ordinarily known as the factor comparison method. It has been described in detail by Benge, Burk, and Hay.[7] In this method, fifteen or twenty "key jobs" are first selected. These are jobs which have present rates not subject to controversy and which are considered by the job evaluation committee to be neither underpaid nor overpaid. These jobs are compared with respect to factors common to all jobs. The factors used in the Benge, Burk, and Hay system are:

Mental Requirements
Skill Requirements
Physical Requirements
Responsibility
Working Conditions

The "key jobs" are ranked in order on each of the factors mentioned, and all appear in each of the lists. The rankings are made independently by approximately ten raters, and are made three times by each rater, with periods of approximately one week intervening between each rating. The present or typical salary of each "key job" is then divided into five

[7] E. J. Benge, S. L. H. Burk, and E. N. Hay, *Manual of Job Evaluation*, fourth edition (Harper and Brothers, 1941).

parts in proportion to the relative importance of the five factors as determined from the ratings. From these steps results a "yardstick" scale based on the key jobs. This scale is then used in evaluating the remainder of the jobs in the company. Each job is compared with the key jobs with respect to each factor and is given a value for that factor. Adding the values thus obtained for each factor results in a total point value for the job. This is the value later used in converting the job evaluation results to a monetary scale to be described later.

3. Comparison of entire job with a predetermined scale

In the operation of this method, a series of job level classifications or brackets is set up by management or by the committee charged with this function. Each job is then assigned to one of the grades. The system is illustrated by such classifications as junior clerk, senior clerk, principal clerk, and so forth, used in the civil service system. In operation, this plan usually results in the grade location of each job being determined largely by the rate which it carried before the job evaluation was made. The plan thus involves a tendency to perpetuate any such inequities as may have existed at the outset.

4. Comparing job elements of each job with a predetermined point scale

This plan, in one of its many variations, is the most widely used system of job evaluation. It is well illustrated by the system used by the National Electrical Manufacturers Association (ordinarily referred to as the NEMA system). The same system is also used by the National Metal Trades Association. This system requires that each job be studied in terms of each of eleven characteristics, and that one of five degrees of each of these characteristics be assigned to the job. The eleven characteristics with the points corresponding to each degree of each are summarized in Table 34.

TABLE 34

Job Characteristics and Point Values Corresponding to Various Degrees of Each Used in the National Metal Trades Association Job Evaluation System

POINTS ASSIGNED TO FACTORS AND KEY TO GRADES

Factors	First degree	Second degree	Third degree	Fourth degree	Fifth degree
Skill					
1. Education	14	28	42	56	70
2. Experience	22	44	66	88	110
3. Initiative and ingenuity	14	28	42	56	70
Effort					
4. Physical demand	10	20	30	40	50
5. Mental or visual demand	5	10	15	20	25
Responsibility					
6. Equipment or process	5	10	15	20	25
7. Material or product	5	10	15	20	25
8. Safety of others	5	10	15	20	25
9. Work of others	5		15		25
Job Conditions					
10. Working conditions	10	20	30	40	50
11. Unavoidable hazards	5	10	15	20	25

The meaning of each "degree" is clarified by specific "degree" definitions, which are an integral part of this system. For example, in the National Electrical Manufacturers Association system, the "degree" definitions for the various amounts of experience required by the job are as follows:

Degree	*Amount of Experience*	*Points*
1	Up to three months	22
2	Over three months up to one year	44
3	Over one year up to three years	66
4	Over three years up to five years	88
5	Over five years	110

Similar "degree definitions" are included for the various degrees of the remaining ten factors or characteristics. This plan has been specifically constructed for the evaluation of shop jobs. A similar system, encompassing a different list of

factors, has been designed for the evaluation of salaried positions; it provides for the rating of salaried jobs on the following factors: education, experience, complexity of duties, supervision received, errors, contacts with others, confidential data, mental or visual demand, working conditions, character of supervision, and scope of supervision. The last two are used only when supervisory duties are involved, which means that all jobs that do not include supervisory duties automatically receive zero points on these two factors.

Under the operation of this kind of plan, each job receives a total number of points (the sum of the points received for each characteristic). These total point values are then used in setting up the monetary wage scale.

Converting Job Evaluation Results to a Wage Scale

Whatever job evaluation system is used, the result is a grading of all jobs according to the principles and assumptions involved in the plan. Ordinarily this grading is entirely completed before any reference is made to the wage or monetary aspect of the work, although in some plans the original evaluations are made in terms of cents rather than points. The experience of most men in this field has shown that original evaluation in terms of cents is inadvisable, because persons on the job evaluation committee are less likely to evaluate jobs on an impartial basis if they think in terms of cents per hour than if they think in terms of points, the cent value of which has not yet been determined.

After the evaluations have been completed, however, it is necessary to convert the results into money in order to build a wage structure. A simple and generally satisfactory method of making this conversion is through the use of key jobs that are common to various plants in the locality. For example, the jobs listed in Table 35 might be found in several plants in a given locality as well as in the plant that is revising its wage structure through a comprehensive job evaluation program. Table 35 also gives for each of the "common jobs" the total

points for the job and the "going rate" for the job in the locality.

The point values of the jobs listed in Table 35 are plotted against the "going rates" for these jobs. The result is illustrated in Figure 120. A line is then drawn so as to show the relationship between points and money. This line may be fitted by the method of least squares[8] if great accuracy is

TABLE 35

TYPICAL JOBS THAT MIGHT BE USED IN CONVERTING JOB EVALUATION RESULTS TO A WAGE STRUCTURE

Jobs Common to Several Plants in the Locality	Job Evaluation Points	Local "Going Rate"
Tool inspector	321	1.45
Engine lathe operator	277	1.20
Mason	280	1.36
Millwright	301	1.45
Pipefitter	289	1.40
Punch press operator	204	1.05
Tool maker	344	1.60
Parts wrapper	97	.60
Janitor	110	.65
Laborer	136	.70

desired. For practical use, however, a line drawn in by inspection of the points is usually sufficiently accurate. From this line, one may read directly the appropriate rate in cents per hour for all other jobs that have been evaluated. Many of these other jobs will be unique to the plant in question and will not be found in other plants in the locality. The use of a chart such as the one illustrated in Figure 120 provides for setting appropriate rates for all jobs so as to form a consistent wage structure.

Inspection of Figure 120 suggests that every slight difference in points results in a corresponding difference in hourly rate. In practice, most companies (and unions) feel that the inherent lack of perfect accuracy in the judgments that

[8] Karl J. Holzinger, *Statistical Methods for Students in Education* (Ginn and Company, New York, 1928), pp. 321 ff.

underly a set of job evaluations makes it desirable to bracket together jobs of approximately the same point value and to consider these jobs as equal in setting up the wage structure. This bracketing results in so-called *labor grades*. The number of labor grades found in specific wage structures varies from

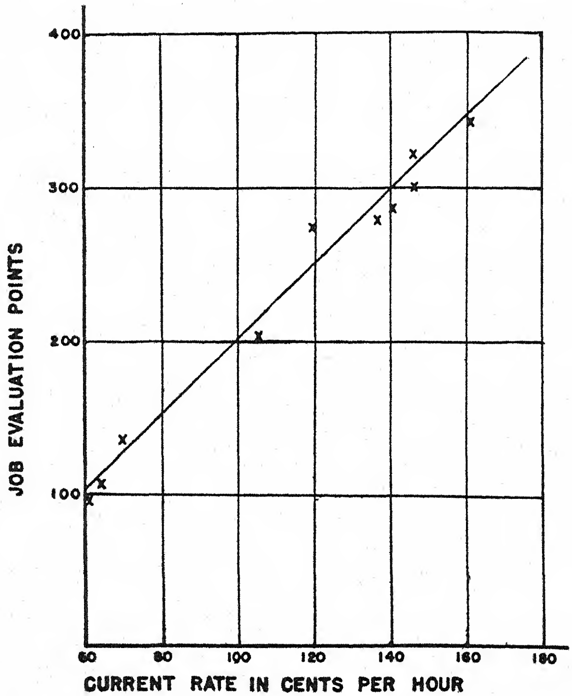


FIG. 120—Job evaluation points of typical "key" jobs plotted against the "going rate" for the jobs in the locality.

around eight or ten to twenty or twenty-five. The tendency of most current union demands in wage contract negotiations is to favor a relatively small number of labor grades.

When the jobs have been bracketed in labor grades, provision is usually made for wage increases within each labor grade, as illustrated in Figure 121. Various procedures have been used in granting wage increases within labor grades, as well as in upgrading employees to higher categories. Some

companies use an automatic acceleration schedule under which specified increases automatically become effective after a specified period of time on the job. This principle is employed most frequently in the lower labor grades and with new employees, but it is also sometimes used at higher levels in the wage structure. A systematic merit-rating program is also used by some plants as a means of identifying employees

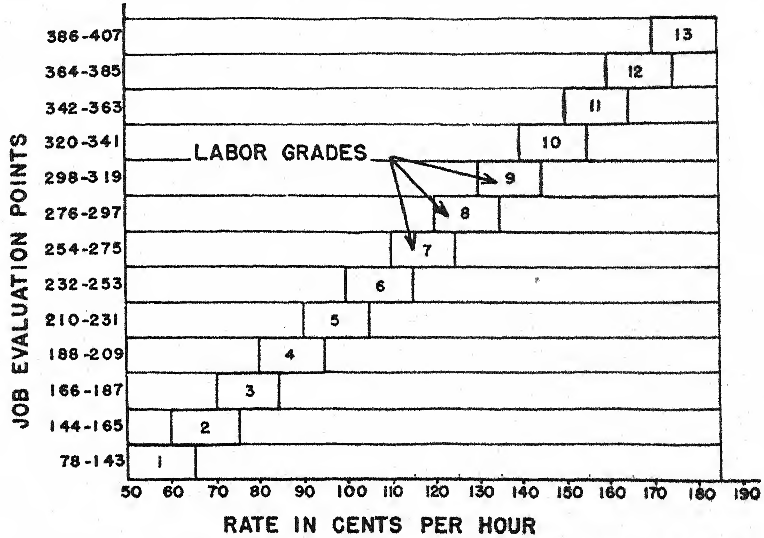


FIG. 121—Grouping of jobs into labor grades, with the range of rate of each grade.

who are eligible for a wage increase under the prevailing wage structure.

The particular method that is followed in making wage increases has often been the subject of bitter controversy in union-management contract negotiations. Unions frequently demand rigid acceptance of seniority as the basic determining factor, and the company as frequently contends that proper managerial responsibility requires the granting of wage increases upon the basis of proved merit rather than upon that of mere time on the job. Some companies have used job qualification tests for the upgrading of employees, not

only within a given labor grade but from one labor grade to another. The Northwest Airlines has in operation such an upgrading program based on paper-and-pencil, as well as performance, tests under terms of a contract with the union concerned, the Air Line Mechanics Association.[9] The use of tests as one of the tools for upgrading offers the dual advantage of assuring members of the bargaining unit that qualified men will be given proper consideration, regardless of their seniority, and of assuring the company that unqualified men will not automatically be given promotions with tenure on the job, irrespective of proved ability.

Weighting of Items in Job Evaluation

Under a job evaluation system, such as the NEMA plan, it is usually considered advisable to assign different weightings to the several characteristics. The weights used are based upon the experience and judgment of men charged with the responsibility of creating and installing a usable system. In the operation of many plans, different weights are given to the various characteristics by allowing a different maximum number of points for the several items.

In the NEMA system the weights assigned are as summarized in Table 36. In actual operation, the functional weightings are determined by the variability of spread of the ratings on each trait.[10] Thus, if the variabilities of the item values were to differ in proportionate value from the maximum values assigned, the actual weightings would not follow the assumed weightings. It is interesting to note, also in Table 36, that in three plants the functional weightings computed from variability measures agree rather closely with the weights based on the assigned maximum values. The variable maximum value technique of assigning weights, therefore,

---

[9] *Trade Test Standard Manual for Written and Practical Tests*, Personnel Department, Training Division, Trade Testing Section, Northwestern Airline Inc., Minneapolis, Minnesota.

[10] See Appendix A, page 501.

would seem to be a fairly effective, as well as a simple, method of giving any predetermined set of weights to the factors involved.

TABLE 36
WEIGHTS OF ELEVEN FACTORS ASSUMED BY MAXIMUM VALUES OF ITEMS IN THE NEMA SYSTEM, AND FUNCTIONAL OPERATION OF ACTUAL WEIGHTINGS IN THREE PLANTS

Item		Percentage Weights Assumed by Maximum Values of NEMA System	Percentage Weights Operating in System in Three Plants		
			Plant A	Plant B	Plant C
Skill	Education	14	18	17	13
	Experience	22	26	27	26
	Initiative and Ingenuity	14	16	15	14
Effort	Physical Demand	10	6	7	8
	Mental or Visual Demand	5	4	3	5
Responsibility	Equipment or Processes	5	5	5	4
	Material or Product	5	5	3	4
	Safety of Others	5	5	4	6
	Work of Others	5	4	7	4
Job Conditions	Working Conditions	10	8	9	8
	Unavoidable Hazards	5	3	3	8

Statistical Studies of Job Evaluation Plans

Any one of the methods of job evaluation results in a reasonably consistent set of evaluations of the various jobs studied. Each is based upon certain assumptions, which differ from plan to plan, and no one system should be considered to be an "exact science," as Lytle[11] has pointed out. As long as *any* systematic plan is followed that provides for

[11] Lytle, *op. cit.*, p. 4.

all jobs to be studied in terms of the *same* assumptions and the *same* other stipulations (such as weights for the factors involved), the result is far more satisfactory than a wage structure that has grown piecemeal, without integration.

The choice of which system to use is not, therefore, a problem of choosing from among several plans, only one of which can be correct, all the others being wrong. There are both practical advantages as well as disadvantages to every system, and any one plan that is installed with care is almost certain to be better than no plan at all. However, the time and expense involved in installing and operating a job evaluation system are such that it is advantageous to use a method that will accomplish the desired result with a minimum of labor, if the results obtained from a simple system are equivalent to those that would result from a more complex one.

Certain basic methods of psychology are ideally adapted to a quantitative study of job evaluation systems. These methods have been applied in a series of studies reported by Lawshe and his collaborators.[12] Since the NEMA point system (or a system basically the same) has been more

---

[12] C. H. Lawshe, Jr., and G. A. Satter, "Studies in Job Evaluation: I. Factor Analyses of Point Ratings for Hourly-Paid Jobs in Three Industrial Plants," *Journal of Applied Psychology*, XXVIII (June, 1944), pp. 189–198.

C. H. Lawshe, Jr., "Studies in Job Evaluation: II. The Adequacy of Abbreviated Point Ratings for Hourly-Paid Jobs in Three Industrial Plants," *Journal of Applied Psychology*, XXIX (June, 1945), pp. 177–184.

C. H. Lawshe, Jr., and A. A. Maleski, "Studies in Job Evaluation: 3. An Analysis of Point Ratings for Salary Paid Jobs in an Industrial Plant," *Journal of Applied Psychology*, XXX (April, 1946), pp. 117–128.

C. H. Lawshe, Jr., and Salvatore L. Alessi, "Studies in Job Evaluation: IV. Analysis of Another Point Rating Scale for Hourly-Paid Jobs and the Adequacy of an Abbreviated Scale," *Journal of Applied Psychology*, XXX (August, 1946), pp. 310–319.

C. H. Lawshe, Jr., and R. F. Wilson, "Studies in Job Evaluation: 5. An Analysis of the Factor Comparison System as It Functions in a Paper Mill," *Journal of Applied Psychology*, XXX (October, 1946), pp. 426–434.

C. H. Lawshe, Jr., and R. F. Wilson, "Studies in Job Evaluation: VI. The Reliability of Two Point Rating Systems," *Journal of Applied Psychology*, XXXI (August, 1947).

frequently used than any other,[13] Lawshe first made certain statistical analyses of the results of several NEMA point system installations to determine whether approximately the same results could have been obtained by the use of a simplified system involving fewer characteristics to be rated.

Using a standard statistical technique,[14] a determination

TABLE 37

CORRELATIONS OBTAINED BETWEEN RATINGS ON SELECTED ITEMS AND TOTAL POINT RATINGS IN THREE PLANTS

Selected Rating Scale Items	Correlation with Total Points Based on 11 Items
Plant A:	
Experience (or learning time)	.96
Experience (or learning time) plus hazards	.97
Experience (or learning time) plus hazards plus education	.98
Plant B:	
Experience (or learning time)	.93
Experience (or learning time) plus initiative	.95
Experience (or learning time) plus initiative plus responsibility for the safety of others	.96
Plant C:	
Experience (or learning time)	.86
Experience (or learning time) plus hazards	.91
Experience (or learning time) plus hazards plus initiative	.93

was made of which single factor, which two factors, which three factors, and so on, give point ratings that correlate highest with the total point ratings. The results obtained

[13] A survey conducted in December, 1944, and reported by Lytle (*op. cit.*, p. 13) showed that 38 of 51 companies surveyed were using job evaluation. Among these 38 companies:

24 were using a point system
8 were using a factor comparison system
5 were using a ranking system
1 was using an unclassified system

[14] The Wherry-Dolittle Technique, one of the basic statistical methods used for this purpose, is described by W. H. Stead, C. L. Shartle, *et. al.*, *Occupational Counseling Techniques* (American Book Company, 1940), pp. 245 ff.

in three different plants using the NEMA system (or a slight modification of it) are summarized in Table 37.

In each case the correlations obtained between the most influential three characteristics and the total points determined from all eleven factors were so high that the contribution of the remaining eight factors was negligible.

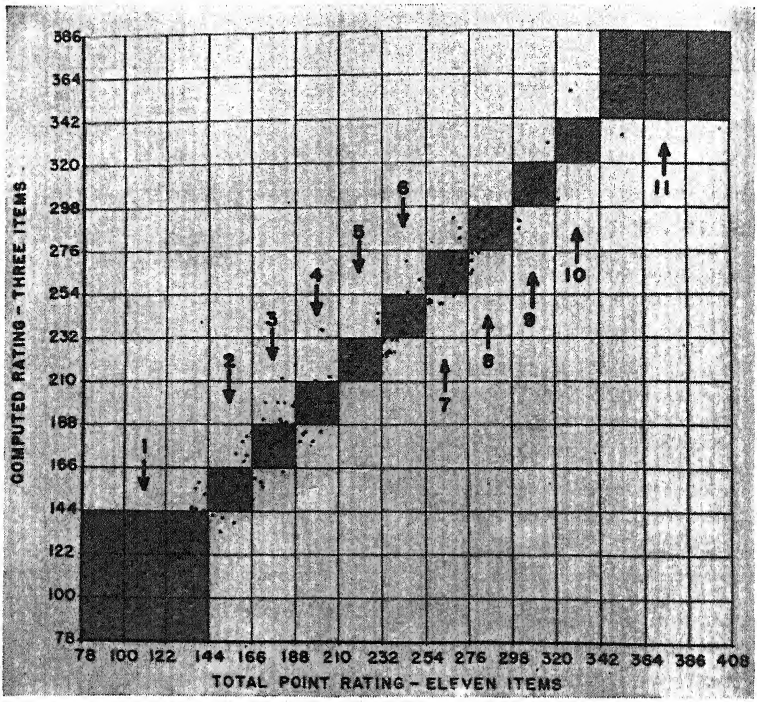


FIG. 122—Graph showing ratings computed from three scale items plotted against total point ratings (all eleven items) for 247 jobs in Plant A. The eleven shaded areas define the labor grades designated by the numbered arrows.

The facts summarized in Table 37 show that experience (or learning time) is the most important single factor in determining the total points for the jobs evaluated in all three of the plants. It will also be noted that *hazards* appear among the most important three factors in two of the plants, as does *initiative.* The results further show that three job

characteristics give results which, for all practical purposes, are statistically equivalent to those obtained from the use of a more extensive system.

Figure 122 is a graphic representation of the results obtained in Plant A showing the accuracy with which the total points for each job can be predicted from a knowledge of the points assigned on only *three* characteristics, namely, experience, education and hazards.

TABLE 38

Discrepancies between Total Point Ratings (Eleven Items) and Ratings Computed from Three Items for Plant A

Points of Deviation	No. of Jobs by Labor Grade Displacement			
	Same Labor Grade	Displaced One Labor Grade	Displaced Two Labor Grades	All Jobs
0–4	68	9		77
5–9	48	27		75
10–14	20	28		48
15–19	11	21		32
20–24	6	2		8
25–29		4		4
30–34		1	1	2
35–39			1	1
Totals	153	92	2	247

The shaded areas in Figure 122 delineate labor grades. From this chart it is apparent that the vast majority of jobs fall in the same labor grade when evaluated on the basis of the stipulated three characteristics as when evaluated on the basis of all eleven factors.

A summary of the effectiveness of the abbreviated three-factor system in Plant A in terms of number of jobs in which the discrepancy in placement between it and the total point rating is one, two, or three labor grades is shown in Table 38.

Table 38 shows that only seven jobs deviate from their

original placement by more than 24 points, the approximate range of points covered by one labor grade in this system.

Lawshe and Maleski[15] have reported a similar study, which made use of the same statistical analysis technique as applied to a set of approximately 400 salary job evaluations. The original system under which the job evaluations were determined was the salary job rating plan devised by the National Metal Trades Association. This plan requires the rating of each job on eleven characteristics, each characteristic to be valued at 5, 6, or 7 degrees. The correlations between results obtained from a reduced number of characteristics and total point values based upon all eleven factors are summarized in Table 39.

TABLE 39

CORRELATIONS OBTAINED BETWEEN RATINGS ON SELECTED ITEMS AND TOTAL POINT RATINGS FOR APPROXIMATELY 400 SALARY JOBS

Selected Scale Items	Correlation with Total Points Based on 11 Items
Experience	.93
Experience plus complexity of duties	.96
Experience plus complexity of duties plus character of supervision	.98

In the results for salary jobs, as in those previously discussed for shop jobs, the experience factor is shown as the most important single characteristic. Also as in the former instance, ratings based on only three characteristics are revealed as equivalent to those based on the larger number of characteristics.

The preceding discussion has been concerned with work on abbreviating the NEMA point systems of job evaluation for shop jobs and salaried jobs. The question may arise as to whether the same pattern of results would be obtained if

[15] Lawshe and Maleski, *op. cit.*

another type of point system was similarly analyzed. In answer to this question, Lawshe and Alessi[16] report a study in which the same type of analysis was made of ratings obtained with a point system differing from the NEMA system in that each job is rated on seven characteristics by means of more and finer categories or degrees for each of the factors, and that the point ratings are translated into "rating factors" by a logarithmic conversion chart for the purpose of assigning monetary equivalents. Table 40 summarizes the results obtained from this analysis.

TABLE 40

CORRELATION COEFFICIENTS BETWEEN RATINGS ON SELECTED SCALE ITEMS AND TOTAL POINT RATINGS OBTAINED WITH A POINT SYSTEM DIFFERING FROM THE NEMA SYSTEM

Selected Items	Correlation with Total Points Based on 7 Items
Responsibility	.92
Responsibility plus manual skill	.97
Responsibility plus manual skill plus working conditions	.98

From the correlations summarized in Table 40 it is apparent that the stipulated three characteristics give results practically identical with those obtained from the more elaborate system.

After observing the results of statistical simplification of point systems of job evaluation, the question may be raised as to whether corresponding steps toward simplification of other kinds of job evaluation plans will be equally effective. A similar type of statistical analysis of the factor comparison system as installed in a paper mill has been reported by Lawshe and Wilson.[17] Here also the results show that a

[16] Lawshe and Alessi, *op. cit.*

[17] C. H. Lawshe, Jr., and R. F. Wilson, "Studies in Job Evaluation: 5. An Analysis of the Factor Comparison System as It Functions in a Paper Mill," *Journal of Applied Psychology*, XXX (October, 1946), pp. 426–434

reduced number of job characteristics gives results that correlate almost perfectly with the total factor evaluations. These correlations are summarized in Table 41.

TABLE 41
CORRELATION COEFFICIENTS BETWEEN RATINGS ON SELECTED ITEMS AND TOTAL POINTS OBTAINED WITH THE FACTOR COMPARISON SYSTEM OF JOB EVALUATION

Selected Items	Correlation with Original Ratings
Skill requirements alone	.94
Skill plus working conditions	.97
Skill plus working conditions plus mental requirements	.99

From the results of the several comprehensive studies reported by Lawshe and his collaborators, it seems rather clear that job evaluation results statistically equivalent to those derived from more elaborate systems can be obtained from a very much abbreviated plan.

Reliability of Job Evaluation

Any form of job evaluation is basically an attempt to measure the relative worth of various jobs by a process of individual, pooled, or committee judgments. The ratings obtained (regardless of the system in use) should have a satisfactory reliability, just as the results of any other type of measurement must have a satisfactory reliability, before usable results can be expected. Other things being equal, if two systems have different reliabilities, the one with the higher reliability should always be chosen.

In order to estimate the reliability of the most commonly used point system (the NEMA plan), and also to estimate the reliability of a simplified point system based upon the work summarized in the preceding sections, a series of reliability studies has been conducted and reported by Lawshe and

Wilson.[18] In this work, forty job descriptions of commonly known jobs were chosen for evaluation. Twenty men experienced in job evaluation made the ratings. Each job was rated by five men using the NEMA plan, and also by five other raters using a simplified plan that included only four items. Each rater made his ratings independently.

TABLE 42

Reliability Data on the NEMA Job Evaluation System and a Simplified Point System Involving Four Characteristics

NEMA System			Simplified System		
(1) *Item*	(2) r	(3) r_5	(4) *Item*	(5) r	(6) r_5
Experience	.82	.96	Learning period	.86	.97
Initiative and ingenuity	.78	.95			
Education	.77	.94	General schooling	.79	.95
Responsibility for safety of others	.54	.85			
Working conditions	.54	.85	Working conditions	.61	.89
Responsibility for work of others	.51	.84			
Physical demand	.47	.82			
Responsibility for equipment or processes	.41	.78			
Responsibility for materials	.40	.77			
Mental or visual demand	.37	.75			
Unavoidable hazards	.34	.72	Job hazards	.51	.84
Total Points	.77	94	Total Points	.89	.98

The four items included in the simplified plan were chosen after factor analyses had been made of several more elaborate systems. These items were selected to cover as fully as possible the basic characteristics revealed by the several factor analyses.

The results of the reliability investigations of the two systems are summarized in Table 42. In this table, the items

[18] C. H. Lawshe, Jr., and R. F. Wilson, "Studies in Job Evaluation: 6. The Reliability of Two Point Rating Systems," *Journal of Applied Psychology*, XXXI (August, 1947).

used in the NEMA system are listed in column (1), in the order of their obtained reliabilities. The obtained reliabilities, listed in column (2), were obtained by computing the coefficient of correlation between the ratings of each rater and those of every other rater for each item in the list. The average of the correlations thus obtained was taken as the best estimate of the reliability of a single rater on the item. In column (3) of Table 42 are given the "stepped up" reliabilities, i.e., the values that should be expected if averages of the ratings of five raters working independently were used consistently.

The four items included in the simplified system are listed in column (4). Each item is located so as to correspond with the approximately equivalent item of the longer scale.

Columns (5) and (6) give, respectively, the obtained and "stepped up" reliabilities of the items of the shorter scale. At the bottom of the table are given the obtained and "stepped up" reliabilities of the total scales.

It will be noted in Table 42 that the items of the shorter scale are all higher in obtained reliability than the corresponding items of the NEMA scale, and this is true whether one considers the obtained reliabilities or the "stepped up" reliabilities. It will be further noted that the reliability of the total scale is higher than the corresponding reliability of the larger scale.

The greater reliability of the shorter scale is probably due to the use of more carefully worded descriptions of the degrees of the various characteristics, rather than to the mere fact that a shortened scale is being used. Familiarity with the scale and experience with its use will not explain the results, because the majority of raters using the NEMA scale had had practical experience with it, whereas *none* of the raters had ever seen the shortened scale before the experiment.

Table 42 also shows that the obtained reliability of the total NEMA point ratings is lower than that obtained for several of the individual factors in this system. In the case

of the simplified system, the total points resulted in a higher reliability than that obtained for any specific factor.

These findings indicate that: (1) the ratings from a shortened system correlate very highly with the ratings of longer systems, (2) the findings are essentially the same for shop jobs and salaried jobs, (3) the effect of shortening the system is the same for point systems and for the factor comparison system, and (4) the ratings obtained from a suggested simplified system have higher reliability than those resulting from a longer system.

In the light of these findings, it is probably advisable to make greater use of systems of job evaluation that are considerably shorter than many of those in current use.

12

Work, Fatigue, and Efficiency

CAREFUL consideration of employee aptitude before placing men upon a job, followed by a thorough program of training employees on the job, will go far toward achieving—but will not guarantee—a satisfactory level of job performance. Other factors besides employee aptitude and training may operate to prevent production from reaching the expected level; and in some cases a proper consideration of these other factors may result in fairly satisfactory production even when little specific attention has been given to employee placement and job training. These other factors have been classified in different ways. They may properly be thought of as related to employee efficiency, job performance, effort expended, or fatigue.

The Nature of Fatigue

At least three conceptions of the term *fatigue* are important to our discussion. These are (1) physiological fatigue, (2) psychological fatigue, and (3) industrial fatigue.

Physiological fatigue

It may be readily demonstrated by simple physiological experiments that a muscle undergoing a simple rhythmic contraction gradually loses its ability to make the contraction. Thus, if one squeezes a coiled spring at intervals of two seconds he will find that his squeezes become less and less powerful until finally, if the task is maintained for sufficient time, only very small changes, or no changes at all, in the spring can be noticed. The energy used in the execution of a task of this

sort, or any muscular task, comes from potential energy that is stored in chemical form in the muscles. As this energy is expended, the muscles become less and less able to perform their task. This reduction in potential energy available in the muscles may be thought of as physiological fatigue. Starling[1] has stated that the phenomena of fatigue, from this point of view, probably depend upon two factors, namely, the consumption of the contractile material or the substances available for the supply of potential energy to this material, and the accumulation of waste products of contraction. These waste products may be thought of as the chemical result of the muscular activity. Among them lactic acid is probably of great importance, because it is known that fatigue may be artificially induced in a muscle by feeding the muscle with a dilute solution of lactic acid.

To the physiologist, then, fatigue is primarily a matter of chemical changes in the muscle itself or the potential energy available for the muscular contraction. It should be kept in mind that this conception, in and of itself, does not involve a feeling of tiredness on the part of the subject, nor an immediate reduction in the amount of work he is able to do, if he is not attempting to do a degree of work that requires more energy than is available at the time.

Psychological fatigue

Entirely aside from the physiological changes occuring in fatigue, it is a matter of common experience that with a repetition of certain tasks one becomes "bored," "uninterested," and in other ways manifests a lessening desire to continue the performance. Or one may simply have feelings of tiredness accompanied by a desire to do something else "for a change." From this point of view, fatigue may be looked upon as the feeling of boredom that often accompanies continued application to any given task.

---

[1] E. H. Starling, *Principles of Human Physiology*, third edition (Lea and Febiger, Philadelphia, 1920), p. 209.

For one speculating upon the subject of fatigue, it is a temptation to look upon these two aspects of the phenomenon —the physiological and the psychological—as different aspects of the same fundamental change. According to this point of view, feelings of boredom and tiredness occur in proportion as physiological changes in the muscle have taken place; and vice versa, actual changes in the physiological composition of the muscles have taken place to the extent that feelings of tiredness or boredom are experienced by the subject.

If this simple relation were true, the problems of fatigue and its elimination would be greatly simplified. Unfortunately, numerous physiological and psychological experiments have shown that no such simple relationship exists. Indeed, it is necessary to look only at one's own experiences to understand that exceptions to such a relationship are by no means uncommon. Anyone who has driven a car a very long distance to a vacation resort knows that interest in reaching the resort is often maintained long after the body, from the physiological viewpoint, would rather sleep. And everyone who has had a routine, uninteresting, manipulative job to do knows that feelings of boredom often become prominent long before an excessive amount of work in the physiological sense has been performed.

We are thus presented with two contrasting, and in many cases unrelated, conceptions of the term fatigue. Because of the great importance of fatigue among industrial employees and the necessity of discovering ways of reducing or eliminating it, the industrial psychologist has been forced to accept a still different definition, a definition that is concerned primarily with the production or output of the employees.

Industrial fatigue

Fatigue is important in industry not to the extent that it involves physiological changes in the muscles, or feelings of boredom on the part of the employees, but, rather, to the extent that it involves on a long or a short term basis a reduc-

tion in the employee's efficiency on the job. Various measures of employee efficiency have been proposed and used for different purposes but, everything considered, the most satisfactory measurement for most purposes is production. Management, therefore, defines fatigue as *whatever changes occur as a result of work that are associated with a decrease in employee production.* This does not mean that management is uninterested in the physiological basis of fatigue or in the boredom and dissatisfaction aspects of the problem. The latter factor indeed has been found more markedly related to employee production than physical or organic conditions. In one series of studies, it was found that "fatigue" usually implies dissatisfaction and that physical conditions varying within ordinary limits have an insignificant effect upon the efficiency of the worker.[2] In another investigation it was found that whatever is done by management to attract the workers' interest or indicate a concern for their welfare decreases "fatigue" and increases output.[3] But in the final analysis, unless physiological or attitudinal changes result eventually in reduced output, on either a long or a short time basis, management cannot justifiably give very much attention to them. The problems of industrial fatigue, therefore, become essentially problems of determining what such factors as conditions of work, posture, hours of work, or nutrition, contribute toward the maintenance of proper production, and what factors result in, or are accompanied by, an undue reduction of employee output. Vernon[4] has defined industrial fatigue as "the sum of the results of activities which show themselves in a diminished capacity for doing work." Since

---

[2] Committee on Work in Industry of the National Research Council, "Fatigue of workers and its relation to Industrial Production" (Reinhold, 1941).

[3] R. H. Flinn, "Fatigue and War Production," *Medical Clinics of North America,* XXVI (1942), pp. 1121–1143.

[4] H. M. Vernon, "Industrial Fatigue in Relation to Atmospheric Conditions," *Physiological Reviews,* VIII (1928), pp. 130–150.

this definition of fatigue has been found to be most satisfactory in industry, it is the definition that we shall follow.

An acceptance of this definition for industrial purposes does not mean that the physiological or psychological definitions are not of definite value. Rather it means that the industrialist is interested in fatigue in these latter senses only to the extent that it results in lowered production. But, just as the physiological and psychological definitions of fatigue do not always agree with each other, so also the industrial definition does not always agree with either of these two. For example, Arps[5] has pointed out that a production curve can be maintained at a fairly high and constant level if incentives are sufficiently strong, even though physiological changes indicate a great or even unwise expenditure of energy. Of course, management is seldom interested in maintaining production by an expenditure of energy that will result in lowered output from a long-time viewpoint. But here again the final criterion as to whether fatigue is really present in the industrial sense is not whether physiological evidences of fatigue can be observed at the time but whether any reduction in output can be noticed, either at the time or in the future, as a result of the activity.

Factors Related to Fatigue and Employee Efficiency

Work methods and motion and time study

Motion and time economy as related to bodily movements, one of the types of job analysis defined on page 28, is concerned with improvement in work method. From ordinary observation and experience it is apparent that different methods of doing a certain job may require different amounts of time and effort.

If an individual has a pile of rocks to move from one location to another he can do the work with less physical

---

[5] G. F. Arps, "A Preliminary Report of 'Work with Knowledge vs. Work without Knowledge of Results,'" *Psychological Review*, XXIV (1917), pp. 449–455.

fatigue by using a wheelbarrow than by carrying the rocks one at a time in his hands. This simple fact indicates the almost unlimited possibilities of fatigue reduction by making use of proper work methods. Gilbreth[6] was able to demonstrate more than thirty years ago that the work of the average bricklayer can be increased from 120 bricks per hour to 350 bricks per hour by following a more economical pattern of movement. This early study with its concrete results pointed the way toward an application of this principle—the lessening of fatigue by improved work methods—in a wide variety of industrial jobs. Among the possible variations in work method are different types of bodily movement that may be required with different methods of operation. The bodily movement aspect of motion study may be considered a legitimate field for psychological investigation[7] although it is ordinarily considered a branch of industrial engineering. Since instruction and laboratory work in this field are usually given by departments of industrial engineering, a thorough coverage of the topic does not come within the scope of this book. The industrial psychologist should, however, be aware of the significance of the field and should supplement his training in psychology with sufficient work in motion and time study to achieve competence in this area.

---

[6] F. B. Gilbreth, "Bricklaying System" (M. C. Clark, New York, 1909).

[7] This field, although basically a part of experimental psychology, has been largely developed in the United States by industrial engineering rather than by psychology, owing in large measure to the fact that when productive industry first began to recognize the importance of motion and time study as a tool of scientific management, American psychologists, with very few exceptions, were interested in other fields and did not develop this area of psycho-technology. The outstanding exception may be found in the work of Frank and Lillian Gilbreth. Mrs. Gilbreth, a professional psychologist, clearly saw the possibilities for applied psychology in this field, and with her husband, a professional construction engineer, published a series of papers and books in which were formulated most of the principles of motion study that are most widely in use today.

In Great Britain the development was somewhat different. The National Institute of Industrial Psychology recognized motion and time study as a legitimate field of psychology and undertook the development of the area that, in America, went by default to industrial engineering.

The field of economy in work methods has been referred to as motion study, as time study, as time-motion study, and, simply, as methods improvement. It would seem that *motion* study unaccompanied by corresponding *time* study is not very meaningful, because one of the most effective ways of comparing different *motions* is in terms of the *time* which they require. A comprehensive discussion of the whole field of methods improvement would be far beyond the scope of this book, because it involves principles of tool improvement, machine design, and other matters that are not psychological in nature.

The principles of motion economy have been divided into three types,[8] namely:

I. Principles of motion economy as related to the use of the human body.
II. Principles of motion economy as related to the management of the work place.
III. Principles of motion economy as related to the design of tools and equipment.

The first of these is the one of most concern to the psychologist. Our discussion will therefore be limited to principles of this type.[9] The first three are as follows:

1. The two hands should begin as well as complete their motions simultaneously.
2. The two hands should not be idle at the same instant except during rest pauses.
3. Motions of the arms should be made in opposite and symmetrical directions, instead of in the same direction, and should be made simultaneously.[10]

---

[8] This classification is taken from R. M. Barnes, *Motion and Time Study*, second edition (John Wiley & Sons, Inc., 1940), p. 146.

[9] The student who is interested in studying these in more detail and also in studying principles falling under Class II and Class III is referred to Barnes, *op. cit.*, Chapters 12–14, and also to M. E. Mundel, *Systematic Motion and Time Study* (Prentice-Hall, Inc., 1947).

[10] These three principles were first formulated by F. B. and L. M. Gilbreth, "A Fourth Dimension for Measuring Skill for Obtaining the One Best Way to do Work," *Society of Industrial Engineers Bulletin*, V (1923), p. 6.

Numerous experimental demonstrations of the time saved by following these principles have been made. Several such experiments are reported by Barnes,[11] typical of which is the job of assembling a bolt and three washers. On this job, in the study reported, it was found that a 53 per cent increase in production resulted from a work rearrangement to permit adherence to the three principles listed above.[12]

The fourth principle reads:

4. Operations should utilize bodily movements that require the least amount of time and effort.

For some operations, finger motions or finger and wrist motions are better than motions involving grosser parts of the body.[13] Controlled investigations have shown, however, that for certain types of activity, wrist motions[14] and forearm motions[15] can be made more rapidly and with less fatigue than finger motions. In another investigation is was found that, at least under certain conditions, wrist and elbow movements are faster than finger or shoulder movements.[16]

The fifth principle:

5. Momentum should be employed to assist the worker whenever possible, and it should be reduced to a minimum if it must be overcome by muscular effort.

This principle was used in F. B. Gilbreth's early studies of bricklaying.[17] Methods were developed which made use of the momentum of the moving brick in forming the mortar into the joints.

---

[11] Barnes, *op. cit.*, pp. 147 ff.

[12] Barnes, *op. cit.*, p. 149.

[13] Barnes, *op. cit.*, p. 161, and H. C. Sampter, *Motion Study* (Pitman Publishing Co., 1941), p. 100.

[14] M. Smith, M. Culpin, and E. Farmer, "A Study of Telegraphers' Cramp," *Industrial Fatigue Research Board*, Report #43 (1927).

[15] R. H. Stetson and J. A. McDill, "Mechanisms of the Different Types of Movement," *Psychological Monographs* XXXII (1923), pp. 37–40.

[16] W. L. Bryan, "On the Development of Voluntary Motor Ability," *American Journal of Psychology*, V (1892), pp. 71 ff.

[17] F. B. Gilbreth, *Motion Study* (D. Van Nostrand Company, 1911).

The sixth principle:

6. Continuous and curved motions should be substituted whenever possible for straight line motions and sudden or sharp changes in direction.

One investigation has shown that 15 to 24 per cent of a work cycle is used in changing hand direction when an abrupt change in direction is made.[18] If the work layout can be arranged to eliminate this kind of lost time, a marked saving in time is effected.

The seventh principle:

7. If possible, *ballistic* movements (the fast, easy motions caused by single contraction of a muscle group, as in swinging a hammer) should be used instead of *controlled* movements (movements requiring a balance between opposing muscle groups, as in writing by the finger-and-thumb method.)

Although there is little experimental evidence on this subject, the experience of industrial engineers indicates that ballistic movements, in comparison with controlled movements, are more powerful, more accurate, faster, and less fatiguing.[19]

The eighth principle:

8. Motions should be arranged to permit a rhythmic, smooth sequence of operations.

Rhythm has been defined in at least two ways—as a repetitive cycle of motions, and as a sequence of accented motions that may or may not be repetitive.[20] The major characteristics of rhythmic action, in the latter sense, are rapid movements through an arc, with a sudden feeling of termination at the end of the movement. Rhythmic motions, therefore, involve more ballistic than controlled motions.

Motion and time study, and the improvement of work methods through such study, require careful attention to the

[18] Barnes, *op. cit.*, p. 164.

[19] Barnes, *op. cit.*, p. 169.

[20] R. H. Stetson, "A Theory of Rhythm and Discrete Succession," *Psychological Review*, XII (1905), pp. 258–270.

layout of the work and the use of the most efficient kind of tools to do the job, in addition to consideration of the motor aspect of the work. Industrial psychologists should become familiar with this general field since they are often called upon to assist in the phase of management with which it is concerned.

Illumination

In a number of investigations, such as those reported on page 200, it has been found that an employee's job performance is related to his visual acuity at the work distance. Visual acuity, however, is dependent not only upon the visual mechanism of the employee but also upon the illumination under which the work is done. The relation between visual acuity and illumination has been plotted by Troland[21] according to data published by König.[22] The results plotted by Troland show that acuity increases over a considerable range of variation of intensity in almost direct proportion to the amount of illumination. These results suggest that employees on certain types of industrial work should be provided with sufficient illumination to insure a satisfactory level of visual acuity. Wetzel[23] has pointed out that gains in visual acuity are very perceptible up to about 40 foot candles.

In Figure 123 are shown the results of an investigation involving operators of key punching machines in a statistical department.[24] The average production and average errors of the operators before the improvement of lighting conditions are plotted, respectively, as 100 units. With the passage of time after the improvement in lighting, the production showed a definite increase, while the errors were markedly reduced.

---

[21] L. T. Troland, "The Principles of Psychophysiology, *Sensation*, II (D. Van Nostrand, New York, 1930), p. 86.

[22] A. König, "Abhangigkeit der Sehschärfe von der Beleuchtungsintensität," Sitzungsber. der Akad. der Wiss., Berlin, Bd. XIII (1897), pp. 559–575.

[23] M. Wetzel, "L'eclairage dans l'industrie," *Recherches et inventions*, VIII (1927), pp. 81–95.

[24] Adapted from Matthew Luckiesch, *Light, Vision, and Seeing* (D. Van Nostrand, 1944), p. 262.

A thorough study of the effects of different systems of lighting upon output and accuracy of employees engaged in fine work has been published by Weston and Taylor.[25] The effect of various degrees of illumination upon the output of employees can readily be studied by means of controlled

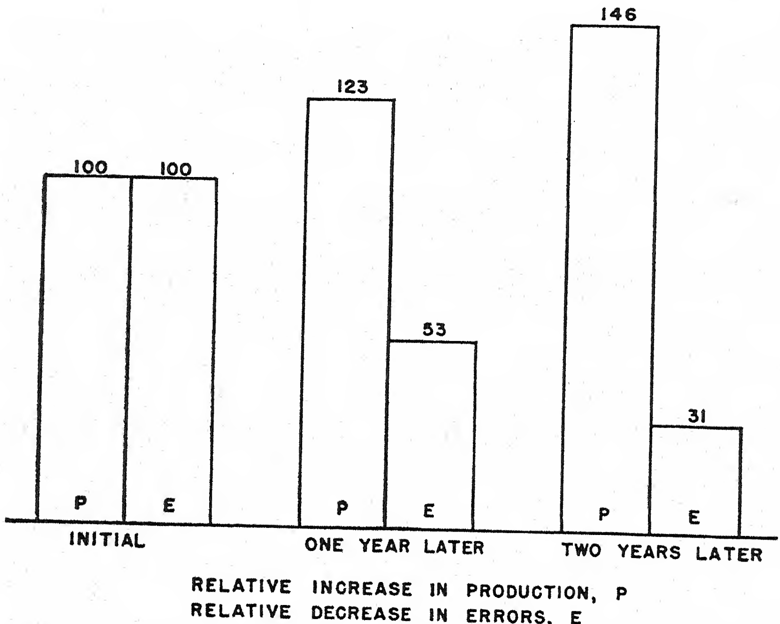


FIG. 123—Production and errors of key punch operators in relation to illumination. (Adapted from Mathew Luckiesch, *Light, Vision and Seeing*, New York, D. Van Nostrand, 1944, p. 262.)

experiments in which illumination is varied and the production measured under varying conditions. Such experiments often make it possible to increase production by an amount that is worth many times the cost of the extra light.

Another factor related to ease of seeing, and therefore also

---

[25] H. C. Weston and A. K. Taylor, "The Effect of Different Systems on Output and Accuracy in Fine Work," *Joint Report of the Industrial Fatigue Research Board and the Illumination Research Committee* (H. M. Stationery Office, London, 1928).

to production, is the contrast between the objects or materials that must be observed and the brightness of the background upon which the objects are seen. The results of an experiment on the effect of brightness contrast upon time required for 100 per cent certainty in seeing are plotted in Figure 124.[26] It will be noted that, for each of the levels of illumina-

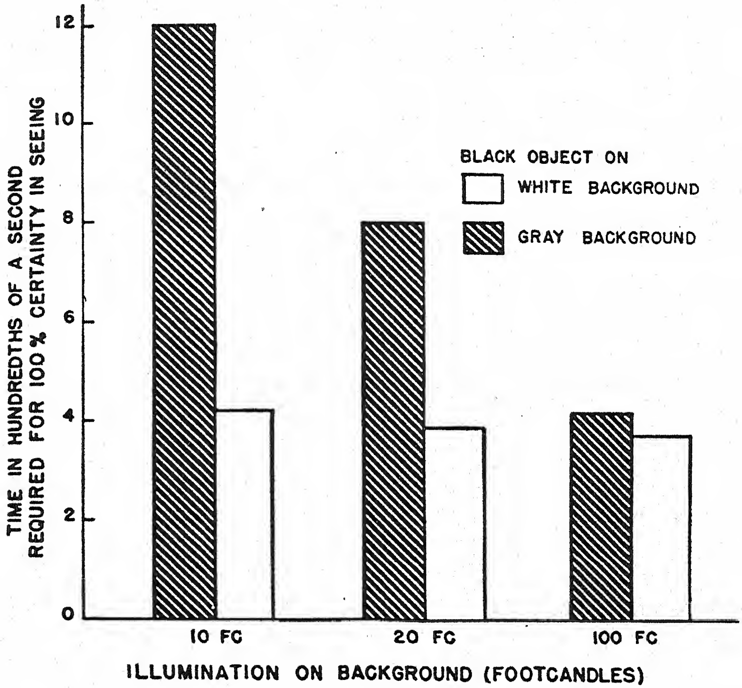


FIG. 124—Relation between brightness contrast and certainty of seeing. (Adapted from Mathew Luckiesch, *Light, Vision, and Seeing*, New York, D. Van Nostrand, 1944, p. 132.)

tion studied, the black object is seen more quickly on the white background than on the gray background.

Visual adaptability to the job

Experiments summarized in Chapter 7 on page 233 show that output of employees on certain types of close work can

[26] Adapted from Luckiesch, *op. cit.*, p. 132.

be definitely increased either (1) by placing on the job only employees who are visually adapted to it or (2) by adapting the employees to the job by means of special occupational spectacles. These experiments may properly be considered as experiments in the reduction of visual fatigue. Indeed, some of the early workers[27] on this subject have considered their work primarily in the field of fatigue. The experiments summarized in Chapter 7 will not be discussed here, but it should be emphasized that the proper allocation of employees according to their visual qualifications is one method of eliminating or markedly reducing visual fatigue. Such reduction often results in an increase or maintenance of satisfactory production.

Temperature and ventilation

The effect of varying temperature upon employee efficiency has not been accurately determined for all types of work, owing partly to the fact that the temperature of the atmosphere, or what is ordinarily known as dry-bulb temperature, seldom varies alone. Almost always when atmospheric temperature varies, other conditions, such as humidity and the amount of radiated heat, do not remain constant. However, in spite of the difficulty of controlling conditions, a few studies have shown that temperature is related to employee efficiency. For example, Vernon, Bedford, and Warner[28] found that when temperature was below 70 degrees, only 3 per cent of the time was lost as a result of employee sickness; when temperatures were from 70 to 79.9 degrees, 4.5 per cent of the time was lost; and when temperatures were 80 degrees or more, 4.9 per cent of time was lost as a result of sickness. These findings indicate that the optimal temperature is below

---

[27] See footnote 29, page 232.

[28] H. M. Vernon, T. Bedford, and C. G. Warner, "A Study of Absenteeism in a Group of Ten Collieries," *Industrial Fatigue Research Board*, Report No. 51 (1928).

70 degrees. In another investigation Yaglogiou[29] found that "the optimum temperature for individuals at rest, or otherwise engaged in light activities in still air and normally clothed" was 64.5 degrees effective temperature. These findings suggest that care should be taken to insure temperature conditions that are not excessively high. Some industries have found that the installation of air-conditioning equipment is a sound investment from the point of view of increased employee productivity.

Although proper ventilation has been found to be more dependent upon circulation of the air than the continual introduction of fresh air, ventilation has also been found to be of considerable importance. Wyatt, Fraser, and Stock[30] report an experiment in which suitably placed fans were operated on alternate working days over a period of six weeks during the summer. Their experiments show that when the air is circulated by the fans the production of the employees is significantly higher than when no provision is made for circulating the air. Bedford[31] has pointed out that the sensations of warmth that one experiences depend upon air temperature, air movement, air humidity, and radiation. All of these factors must be considered in insuring optimum conditions for work.

Noise

A great deal has been written concerning the effect of noise upon employee efficiency and fatigue. Often the elimination of noise has been recommended on the grounds that employee efficiency will be greatly improved. Indeed, several experi-

---

[29] C. P. Yaglogiou, "Modern Ventilation Principles and Their Application to Sedentary and Industrial Life," *Journal of Personnel Research*, III (1925), pp. 375–396.

[30] S. Wyatt, F. A. Fraser, and F. G. L. Stock, "Fan Ventilation in a Humid Weaving Shed," *Industrial Fatigue Research Board*, Report No. 37 (London, 1926).

[31] T. Bedford, "Requirements for Satisfactory Ventilation and Heating," *The Human Factor*, X (London, 1936), pp. 246–254.

ments, such as the one reported by Scheidt,[32] have shown that noise-proofing the workroom or putting employees in individual booths will improve the quality of their work. For the practical industrialist, however, the elimination of noise often is extremely difficult or even impossible.

One can hardly imagine a steel mill in which noise has been eliminated or even much reduced. In steel, as in heavy industry in general, noise can seldom be avoided. Therefore, even though it is true that fatigue may be lessened by reducing noise, we will not suggest such reduction as a practical method of eliminating fatigue in all situations. Too often it is impractical to reduce the noise. It should be kept in mind, however, that when a high noise level is present, employee efficiency may be affected and that it is often possible to recapture some of the lost efficiency by introducing certain other controls in the form of nourishment, rest pauses, or certain favorable work methods.

Music

A number of surveys have been conducted and reported to show how employees feel toward music while they work. Prominent among these has been the work of Kerr[33] and Kirkpatrick.[34] In one of the investigations it was found that approximately two thirds of the employees stated that they prefer to work with music, whereas nearly all of the remaining one third were indifferent. Practically none reported that they prefer to work without music. Of those who prefer to work with music, only 14 per cent preferred to have music played all the time. Beckett[35] concludes from a survey of

---

[32] V. P. Scheidt, "The Effect of Various Modifications on the Worker," *Kalenda* XVI (1937), pp. 3–4.

[33] W. A. Kerr, "Psychological Effects of Music as Reported by 162 Defense Trainees," and "Factor Analysis of 229 Electrical Workers' Beliefs in the Effects of Music," *Psychological Record,* V (1942), pp. 205–212 and 213–221.

[34] F. H. Kirkpatrick, "Music in Industry," *Journal of Applied Psychology,* XXVII (1943), pp. 258–274.

[35] W. Beckett, *Music in War Plants* (War Production Board, Washington, D. C., 1943).

78 plants for the War Production Board that music is universally liked by the workers.

Music may, however, be liked by employees even when it has no effect upon the quantity or quality of their work, or even, conceivably, when it has an unfavorable effect upon the efficiency of the workers. A number of investigations have therefore been conducted to determine the effect of music upon worker efficiency. Wyatt and Langdon[36] found the production of a small group of women on light repetitive work to be increased from 6 to 11 per cent by varying amounts, distributions, and types of music. Humes[37] concluded from another study that music tends to lower the scrappage rate. Kerr,[38] in a series of studies, found that music increases production up to 6 per cent on operations paid on an incentive plan, and up to nearly 12 per cent on hourly-paid jobs. A recent study by Smith[39] has further demonstrated the favorable effect of music upon the production when the work is repetitive in nature. No significant differences were found between accident experience in plants with industrial music and in those without it.

Nourishment

The energy one expends in work comes from the food he eats. Without food, or with food that is insufficient either in quality or amount, one soon loses the ability to withstand physical exertion. Too often the relation between the price of food and its nourishment or energy value is none too high. Some industries have found that a program of employee

---

[36] S. Wyatt and J. N. Langdon, "Fatigue and Boredom in Repetitive Work," Report No. 77 (Industrial Health Research Board, London, 1937), pp. 30–42.

[37] J. F. Humes, "The Effect of Occupational Music on Scrappage in the Manufacture of Radio Tubes," *Journal of Applied Psychology*, XXV (1942), pp. 573–587.

[38] W. A. Kerr, "Effects of Music on Factory Production," *Applied Psychology Monograph* 5 (1945).

[39] H. C. Smith, "Music in Relation to Employee Attitude, Piecework Production and Industrial Accident," National Industrial Conference Board, *Studies in Personnel Policy*, No. 78 (1947), pp. 21–28.

education dealing with types, caloric values, vitamin and mineral content of food, and, in general, how to get the most value for one's food dollar, has been of value in improving the diet of employees. A few industries have found that giving extra meals on the job is worth while in terms of increased employee efficiency. Haggard[40] reports that when two extra meals were given, output rose 10 per cent and the workers reported feeling less tired in spite of their greater output. It has also been reported[41] that drinking a one per cent salt solution is effective in reducing industrial fatigue.

Hours of work

During the past fifty years the trend toward a reduced number of hours in the working week as well as a shortened working day has been fairly constant. This change has been due partly to the feeling of labor that life should not be "all work and no play," but it has been also due in no small measure to an increased recognition by management that employees do not reach their greatest efficiency when the number of working hours is excessive. Many experiments have been conducted to determine the relation between the hours of work and employee production. The reduction in hours of work has been largely justified by these experiments. Miles and Angles[42] report that when the hours of work in a plant manufacturing boxes were reduced from 48 to 36 per week, the average hourly output increased from 793.5 to 834.0 units of production. In another experiment[43] dealing with women engaged in the work of turning fuse bodies, a 68 per cent increase of hourly output and an increase of 16

---

[40] W. W. Haggard, "Work and Fatigue," *Mechanical Engineering*, LVIII (1936), pp. 298–301.

[41] A. B. Dill, A. V. Bock, H. T. Edwards, and P. H. Kennedy, "Industrial Fatigue," *Journal of Industrial Hygiene*, XVIII (1936), pp. 417–431.

[42] G. H. Miles and A. Angles, "The Influence of Short Time on Speed of Production, II," *Journal of the National Institute of Industrial Psychology*, II (1925), pp. 300–302.

[43] C. S. Myers, *Mind and Work* (Putnam, London, 1921).

per cent in total production followed a reduction from 66 to 48.6 in the number of hours worked per week. In another study[44] it was found that an increase in working time above 60 hours per week is accompanied by increase in time lost due to sickness, injury, and absence without permission. The increase in working time is also accompanied by a decrease in both hourly and weekly output and an increase in labor turnover. Murray[45] reports that a reduction of the working hours to less than 48 per week is followed by increases in both hourly and weekly production as well as by decreases in absenteeism and turnover. The general conclusion suggested by these and similar studies is that employee fatigue is lessened and output correspondingly increased if the number of working hours per week does not much exceed 40.

Of course, during a period of emergency production, it is not always possible to achieve maximum efficiency by means of a short working week. At such times it is often necessary to get out maximum production even though by so doing maximum efficiency, from the long-time viewpoint, may be sacrified to some extent. But even during a period of maximum production the question may be raised as to whether real gain can be achieved by excessive lengthening of work hours. Ivy[46] has suggested that from the physiological viewpoint we will sacrifice rather than gain, even in terms of short-term production, by increasing the hours of work beyond 56 hours per week for men or 48 hours per week for women.

Rest pauses

A great many investigations on the effect of rest pauses have been conducted. These studies have considered employees on both heavy and light work, and employees paid under

---

[44] Anon., "Hours of Work, Lost Time, and Labour Wastage," *Emergency Report No. 2* (Industrial Health Research Board, London, 1942).

[45] H. M. L. Murray, "Bases of Worker Efficiency," *Personnel Journal*, XXI (1942), pp. 131–145.

[46] A. C. Ivy, "The Physiology of Work" (Fourth Annual Congress on Industrial Health, Chicago, January, 1942).

a straight hourly rate as well as employees paid under a wage incentive plan. In this field, the work of Vernon[47] and his associates is outstanding. Their studies have shown that rest pauses usually lessen fatigue and increase production in spite of the fact that the pauses result in some reduction of the actual working time. Typical curves showing the production for a group of employees before and after the introduction of rest periods are shown in Figure 125. These curves are based upon the work of Farmer and Bevington.[48]

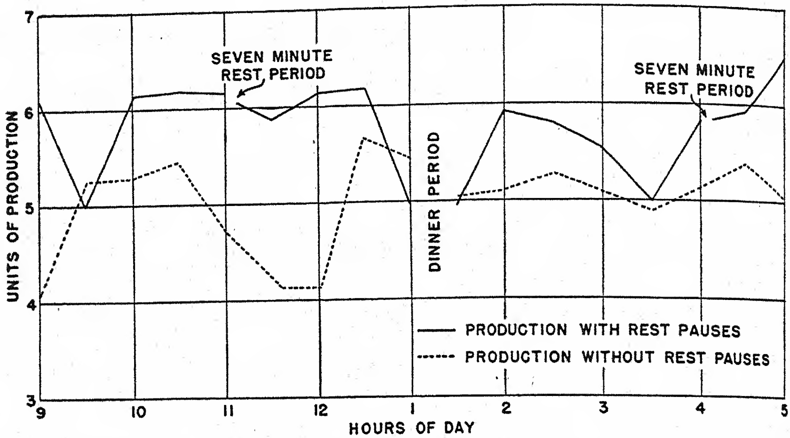


FIG. 125—Effect of rest pauses on production for a typical industrial job.

In another study, Miles and Skilbeck[49] found that two 15-minute change-of-work periods resulted in a 14.2 per cent efficiency increase.

It is impractical to generalize as to how long rest periods should be for any given plant, or how frequently they should be allowed during the day. Both the duration and the

[47] H. M. Vernon, *Industrial Fatigue and Efficiency* (Dutton, New York, 1921).

[48] E. Farmer and S. M. Bevington, "An Experiment in the Introduction of Rest-Pauses," *Journal of the National Institute of Industrial Psychology*, I (1922), pp. 89–92.

[49] G. H. Miles and O. Skilbeck, "An Experiment on Change of Work," *Occupational Psychology, London*, XVIII (1944), pp. 192–195.

number of rest periods necessarily vary with such factors as the nature of the work, the effect of a temporary stoppage upon the flow of production, and the availability of suitable rest rooms or other places in which the employees can relax. The duration of rest periods that have proved beneficial has varied from two minutes to fifteen minutes under different conditions, and the number of rest periods during the day has varied from one to five or even more. An excellent summary of representative studies in this field has been given by Viteles[50] and will not be repeated here. Anyone who wishes to consider the introduction of rest pauses for an industrial job should make a careful study of the literature to find an experiment in which rest pauses have been introduced on a similar job. It is sufficient to say here that such a study will almost certainly show the effect of rest pauses on a similar job and will point the way toward the type of rest pause or pauses that may profitably be introduced. In practically every investigation of this subject, some form or type of rest pause or pauses has been found to be worth while both from the point of view of employee comfort as well as that of increased production.

In considering the introduction of systematic rest pauses, it should not be assumed that if no rest pauses are formally allowed the employees will not rest. Several investigations have shown that if formal rest pauses are not allowed, unauthorized rests will be taken by the employees to suit their own convenience.[51] In some cases, such unauthorized rests are as effective in increasing production and allaying fatigue as rest pauses that are officially and formally allowed by management. But in other cases, employees who decide upon their own rest periods may do so in an ineffective manner, or may inconvenience the production schedule by

---

[50] M. S. Viteles, *Industrial Psychology* (W. W. Norton, New York, 1929) pp. 470–482.

[51] R. B. Hersey, "Rests—Authorized and Unauthorized," *Journal of Personnel Research*, IV (1925), pp. 37–45.

resting at the wrong time. McGehee and Owen[52] found that the introduction of two short rest pauses during the day reduced unauthorized rest pauses and increased the speed of work. In general, it is a more satisfactory procedure for management to decide upon the proper amount and distribution of rest periods and then formally to allow those rests. Just as a person left to learn a job by himself may, and often does, adopt an inefficient and cumbersome way of performing the job (see page 310), so an employee left to decide upon optimal rest pauses for himself will not always adopt rest pauses that are of maximum value either to himself or to departmental production.

Financial incentives

People seldom do something for nothing. No one—whether employee or supervisor, executive or clerk, teacher or student—will achieve a very high level of performance on his job unless he is rewarded for doing so. Rewards may be financial payments in proportion to the actual performance on the job, or they may be nonfinancial returns that are desired by the person concerned. With the rise of modern scientific management, industry has devoted a great deal of attention to the subject of financial incentives. Several basic systems have been worked out that operate to pay the employee in accordance with the quantity or quality of work done. The effect of changing a wage-payment plan from a straight hourly rate to some form of incentive system has been shown in many investigations and is well illustrated in a study reported by Kitson (see page 291), in which a marked increase in employee production was found to follow the installation of an incentive plan of wage payment.

It should, of course, be kept in mind that too great a reward for high production may result in an unwise expendi-

[52] W. McGehee and E. B. Owen, "Authorized and Unauthorized Rest Pauses in Clerical Work," *Journal of Applied Psychology*, XXIV (1940), pp. 605–614.

ture of energy by the employees. An incentive plan that encourages employees to overdraw their physiological bank account of tomorrow because of large financial payment today is not a sound plan from the point of view of either employees or management. But in many cases, as illustrated by Kitson's experiment, low production is due primarily to a lack of reward for high production. After the installation of an incentive plan, the production was practically doubled and no unwise expenditure of physiological energy, either at the time or later, was indicated. The wage incentive plans used by industry may be administered so as to encourage high production up to a certain level and discourage production above that level if such production is likely to result in too great an expenditure of physiological energy.

Nonfinancial incentives

Men want, and will work for, many things besides money. On pages 458 and 459 are listed a number of nonfinancial returns that men want from their jobs. It is becoming increasingly apparent to management that without these "other things" no amount of financial reward, in and of itself, will result in complete job satisfaction. Nor will financial payment alone produce satisfactory job performance. Miles[53] has stressed the importance of nonfinancial incentives, and Lee[54] has quite correctly pointed out that financial incentives are not, by themselves, sufficient to produce "the will to work." The major things that employees want from their jobs, other than money, will be discussed in some detail in Chapter 14. For the present we wish only to call attention to the fact that, as Dickinson[55] has suggested, nonfinancial incentives will reduce fatigue and increase production.

---

[53] G. H. Miles, "Effectiveness of Labour Incentives," *The Human Factor*, VI (1932), pp. 53–58.

[54] C. A. Lee, "Some Notes on Incentives in Industry," *The Human Factor*, VI (1932), pp. 180–186.

[55] Z. C. Dickinson, "American Trends in Industrial Incentives," *Occupational Psychology*, XII (1938), pp. 17–29.

Attitude and morale

The importance of employee morale is so great that a later chapter will be devoted to this subject. Dissatisfied or disgruntled employees readily become bored with their jobs, "tire" of doing their work, and in many other ways show evidence of feelings and behavior that result in lowered output The so-called "fatigue" of such employees is seldom due to excessive expenditure of physiological energy. Rather, it is due primarily to a basic dissatisfaction which, interestingly enough, can seldom or perhaps never be remedied by wage increases alone. Factors that are related to employee morale and suggestions for improving it are discussed in detail in Chapter 14.

Mental abilities and other personality traits

The "born salesman" quickly becomes dissatisfied if he is forced to work by himself on a job that permits little or no contact with other people. The inventive or ingenious employee soon becomes bored with a job that involves only routine manipulative assembly. It has been emphasized in Chapter 1 that the proper placement of employees consists basically in assigning each employee to a job that is neither above nor below his basic capacity and is in reasonable conformance to his basic interests. Definite evidence shows that when employees are not so placed they neither remain very long on the job (see page 2) nor perform the job very well while they remain on it (see page 97). Although such misplaced employees probably do not become unduly fatigued in the physiological sense, they do become fatigued in the psychological sense; and the results of this psychological or mental fatigue are just as serious in terms of reduced production or lack of attention to the job as would be the case if the employees actually became physically exhausted. Wyatt, Langdon, and Stock[56] have emphasized the fact that boredom

[56] S. Wyatt, J. N. Langdon, and F G L. Stock, "Fatigue and Boredom in Repetitive Work," Report No. 77 (Industrial Health Research Board, London, 1937).

is associated with intelligence, inability to mechanize certain simple processes, and the desire for creative work. A typically bored worker is one who is on a job that is below his natural level of capacity and ability. To provide for the training of such employees for a job that involves responsibility commensurate with their capacity, and to arrange for their transfer to the new job when the training has been complete, is a clearcut responsibility of management. Such a policy of upgrading employees who have the capacity pays dividends both in terms of improved production of employees on the routine manipulative jobs and in terms of furnishing a valuable source of employees for supervisory and other responsible positions.

13

Accidents and Safety

ACCIDENTS do not just happen. In dealing with them nothing is gained, except a temporary dodging of responsibility, by ascribing them to bad luck or chance. The concept of luck is used in explaining human behavior only when the really causative factors are unknown or when, for some reason, we have been unable to exercise adequate control over known factors. As those factors that are related to industrial accidents become more and more clear in the light of statistical investigation, the use of the concept of luck or chance becomes less and less necessary. In the light of such causative factors as are now known definitely to exist, Heinrich[1] has estimated that around 98 per cent of industrial accidents are preventable. Of these, around 90 per cent involve such things as faulty inspection, inability of the employee, poor discipline, lack of concentration, unsafe practice, and mental or physical unfitness for the job. These factors may be eliminated or greatly reduced by adequate supervision and proper job placement by employers. The importance of what might be called the psychological or personal factors in causing accidents is firmly attested by Heinrich's conclusion that only around 10 per cent of industrial accidents are due to distinctly physical causes such as faulty equipment or bad building conditions.

General Considerations

Psychological versus physical conditions as causes of accidents

It is the purpose of this chapter to point out a number of essentially psychological factors that have been found by

[1] H. W. Heinrich, *Industrial Accident Prevention* (McGraw-Hill, 1931).

experimental and statistical investigation to be definitely related to industrial accidents. It is not meant, however, to imply that all accidents are related to, or can be explained by, the psychological factors involved. The importance of such external conditions as machinery safeguards and working conditions is fully attested by the favorable safety records that have so often followed the installation of a program that deals primarily with these factors. The present discussion is limited primarily to the psychological aspects of safety, not because we fail to recognize the importance of the other aspects, but because this discussion is primarily intended to bring out the mental and personal features of a safety program.

Lost-time accidents, home cases, and first-aid cases

Any investigation intended to discover factors related to accidents must consider the severity of the accidents. One classification of accident experience makes use of three categories, namely:

1. Lost-time accidents—those in which the employee loses time from the plant in addition to the day, shift, or turn in which the accident occurred. Lost-time accidents involve the payment of industrial compensation.

2. Home cases—those in which the employee loses the remainder of the shift or turn on which the accident occurred. Accidents in this category do not involve the payment of industrial compensation.

3. First-aid cases—those in which the employee receives first-aid attention at the plant hospital or a first-aid station and then returns to his job.

The reason for classifying accidents according to a breakdown of this type is that the factors related to frequency of minor accidents are not all related to that of more serious accidents. This fact is definitely revealed in Table 43, which shows the relative frequency of lost-time accidents and first-aid accidents for eleven departments of a steel mill for the year 1940.

For both types of accidents, the figures are given in terms of accidents per million man hours. A large difference in the relative frequency of accidents of the two types will be noted, the first-aid accidents far exceeding in number the lost-time accidents. A rather great difference in the order of importance or frequency of these accidents in the different departments will also be noted. For example, Department 7, which stands at the top in terms of first-aid accidents with 930 such accidents per million man hours for the year in question, is third from the top in lost-time accidents with 4.18 per million

TABLE 43

ACCIDENT DATA FROM A STEEL MILL FOR ONE FISCAL YEAR

Department	*Lost-Time Accidents per Million Hours*	*First-Aid Accidents per Million Hours*
1	5.91	433
2	3.30	660
3	2.36	574
4	2.34	485
5	8.37	496
6	2.62	392
7	4.18	930
8	2.45	725
9	2.65	504
10	2.26	534
11	.50	242

man hours. Department 5, which stands at the top in lost-time accidents with 8.37 per million man hours, is seventh from the top in first-aid accidents with 496 per million man hours. The rank order coefficient of correlation between frequency of these two types of accidents for the eleven departments studied was only .21. The conclusion indicated by this low correlation is that only a very slight relationship exists between the hazard of a given department with respect to the possibility of lost-time accidents and the hazard of that same department with respect to more minor accidents. Thus, the factors that result in a large number of first-aid or minor accidents do not necessarily result in a correspondingly large number of lost-time accidents, and vice versa. Throughout the discussion of this subject, therefore, we shall

identify in the case of each statistical summary whether the data refer to lost-time or minor accidents.

The importance of reducing lost-time accidents is indicated by the fact that the plant from which a number of the investigations reported in this chapter were obtained has estimated that a lost-time accident involves a minimum cost of $2000 to the company. First-aid injuries, on the other hand, are considerably less expensive. However, no actual figures on the exact cost of these latter injuries are available.

Accident proneness

When industrial safety first began to be given serious consideration by management, opinion differed considerably as to the relative importance of factors within the individual and working conditions as causes of accidents. It is obvious that a safety program within a plant must depend upon specific knowledge of whether these factors are both operating, and, if they are, of what their relative importance is for the plant in question. The concept of accident proneness is concerned with whether each employee tends to retain a given accident rate in comparison with other employees regardless of changes in the general physical condition of the machinery and working conditions within the plant. To the extent that each employee does tend to retain his same relative accident rate from one time to another, we may conclude that personal factors, peculiar to that individual, are affecting his accident record. This situation has been called *accident proneness*. Accident proneness is a descriptive rather than an explanatory term. When an individual employee is found consistently to experience more accidents than the average employee, he may properly be classified as an accident-prone employee. Such classification does not, however, explain *why* he is accident prone; an explanation would require identification of the *causes* of his accident proneness.

The relative importance of working conditions and accident proneness as causative factors varies from one

plant to another. In plants where the machinery used is unavoidably dangerous in itself, the former is of greater importance. In plants where the working conditions in and of themselves are not hazardous, the latter factor is of greater importance. When individual or personal factors are causing the trouble, it is clear that an adequate safety program must make provision for the individual handling of the high accident rate or accident-prone employees. It is, therefore, of vital importance for the safety director of a plant to know the extent to which his own employees are affected by accident-proneness factors.

In nearly every investigation of industrial accidents, accident proneness has been found to be a factor, and in some cases, a vitally important factor. The work of Greenwood and Woods,[2] Marbe,[3] and others whose work has been summarized by Viteles[4] tends to support the statement that "accidents do not distribute themselves by chance, but they happen frequently to some men and infrequently to others as a logical result of a combination of circumstances."[5] Several more recent investigations substantiate this general conclusion. The work of Slocombe,[6] for example, revealed a correlation of .51 between accidents for two successive years. Slocombe concluded, among other things, that the accident-prone men, as identified from their record for a given period of time, will subsequently experience more accidents than the average in 76 per cent of the cases. It is not our intention here to resummarize the excellent digest of the literature on this subject already published by Viteles. Our purpose,

---

[2] M. Greenwood and H. M. Woods, "The Incidence of Industrial Accidents, With Special Reference to Multiple Accidents," *Industrial Fatigue Research Board*, No. 4 (1919).

[3] K. Marbe, *Praktische Psychologie der Unfälle und Betriebschäden* (Munchen, 1926).

[4] M. S. Viteles, *Industrial Psychology* (W. W. Norton, 1932).

[5] *Ibid.*, page 334.

[6] C. S. Slocombe, "Consistency of Operating Efficiency," *Personnel Journal*, VIII (1930), pp. 413–414.

TABLE 44

Frequency of Hospital Visits for Two Consecutive Years in Eleven Departments of a Steel Mill

Department	Freq. 1938	Freq. 1939	Freq. 38–39	Average Number of Hospital Visits in 1939 for Men With 0, 1, 2, etc., Visits in 1938									
				Number of Visits in 1938									
				0	1	2	3	4	5	6	7	8	9
1	.66	.85	.75	.61	1.05	1.54	1.14	1.64	2.20	5.00	2.50	....	....
2	1.13	1.31	1.22	.68	1.31	1.60	2.35	3.63	3.27	3.00	4.33	....	3.50
3	.81	1.03	.96	.63	1.14	1.56	2.16	2.56	4.50	3.00	1.00	....	7.00
4	.89	.81	.94	.55	.87	1.06	1.54	1.86	1.78	2.66	4.25	5.00	....
5	1.01	.99	1.00	.68	.95	1.00	1.94	2.23	1.60	3.33	5.00	....	....
6	1.20	1.27	1.23	.74	1.24	1.56	1.98	1.82	3.63	3.63	5.00	....	7.00
7	1.12	1.26	1.19	.81	1.15	1.45	2.00	2.29	3.44	3.21	11.00	4.00	5.00
8	1.32	1.21	1.26	.64	1.04	1.35	2.34	2.28	2.61	3.36	4.00	6.50	4.70
9	.71	1.05	.88	1.03	1.11	1.43	1.48	3.15	4.50	7.00	....	....	....
10	.91	1.16	1.03	.53	1.26	2.05	2.41	3.18	3.50	4.25	4.50	7.00	8.00
11	.58	.88	.73	.71	1.22	1.47	1.14	3.00	5.50	....	....	....	....
Average for All Depts....	.94	1.07	1.01	.69	1.12	1.46	1.86	1.91	3.32	3.84	4.62	4.91	5.14

rather, is to present certain statistics not previously published that have a direct bearing on this topic. These statistics both strengthen the case for the concept of accident proneness in industrial accidents and answer certain questions which previous investigations have not particularly considered.

Table 44 summarizes certain facts pertaining to hospital visits from the employees of eleven departments of a sheet and

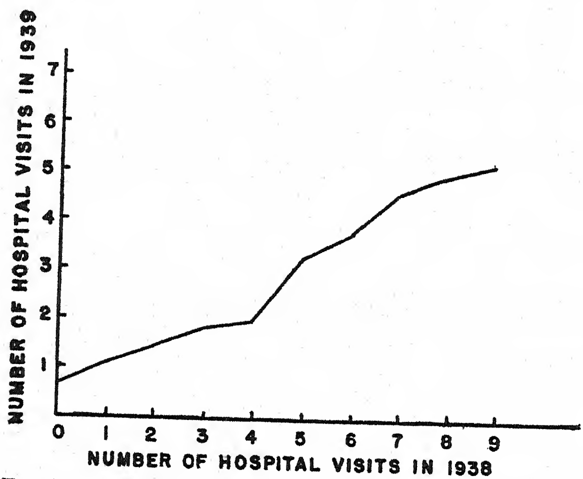


FIG. 126—Relation between number of hospital visits for two successive years among approximately 9,000 steel workers.

tin mill employing around nine thousand employees for the years 1938 and 1939. The second, third, and fourth columns reveal respectively the average number of hospital visits per man for the year 1938, the year 1939, and for the two years combined, for each of the eleven departments as well as for the mill as a whole (in the bottom row). In the remaining columns of this table the average hospital visits per man for the year 1939 are given for the employees who had previously been classified according to the number of hospital visits experienced in 1938. Thus, for Department 1, the individuals with no hospital visits during 1938 averaged .61 hospital

visits in 1939, whereas in this same department the employees with seven visits in 1938 averaged 2.50 visits in 1939. Although some inversions of the data from one category to the next occur, particularly in departments where the total number of men was not large, the general trend in all departments is toward an increase from left to right. When attention is turned to the data summarizing the mill as a whole, which are tabulated in the lower row, this trend is seen to be still more pronounced. For the mill as a whole, no inversions whatever are found in the tendency.

These data for the entire mill are shown graphically in Figure 126. In this figure the number of hospital visits for 1938 is plotted along the base line or abscissa and the corresponding number of hospital visits for the following year of 1939 along the ordinate. From these data we may confidently state that the average employee tends to retain approximately his same relative position with regard to hospital visits for any two successive years.

Accident proneness versus the job in relation to accidents

The question may be raised as to why those employees who had nine hospital visits during the first year did not, on the average, have nine such visits during the second year instead of 5.14 visits as revealed by the data. If factors within the individual, or accident proneness, were the sole cause of hospital visits, then the employees with nine hospital visits during 1938 should also have had on the average nine such visits during 1939. The discrepancy between the nine visits experienced by employees in this group during 1938 and the 5.14 average visits experienced by these same employees in 1939 may be taken as indicative of the relative relation of accident-proneness factors versus external factors to hospital visits in the mill in question. In other words, we may say roughly that of every nine hospital visits made by the most accident-prone group, slightly more than five are essentially related to the accident proneness of these individuals and the

remaining four cannot be related so clearly to individual factors within the employees. To prove the importance of the accident-proneness hypothesis, it is not necessary that the high-accident group for any year retain the same level of accidents for a subsequent year, but only that this group account on the subsequent year for a number of accidents significantly greater than the average of the plant. As revealed in Table 44, this situation is clearly the case for the mill in which these statistics were gathered.

An objection may be raised to the results given above as evidence of the fact that an individual or personal factor operates in the causation of accidents. This objection is that the job to which each employee is assigned may involve a specific accident hazard which, remaining the same or relatively the same from one year to another, results in a consistency of accidents for each employee from year to year, a consistency due more to the accident hazard of a given job than to any factors within the individual. It must indeed be admitted that each job has a relatively constant accident hazard and that this hazard clearly varies by rather marked amounts from one job to another. This fact is brought out in Table 45, which gives the average number of hospital visits per year for employees on eleven specific jobs in a steel mill. It will be noted that the average number of hospital visits varies from 3.55 in the case of cranemen down to .47 in the case of roll turners. It might reasonably be expected that this ratio of 7 to 1 found among this sampling of eleven jobs would be still larger if a larger sampling of jobs had been included in this survey. The table will suffice, however, to set forth clearly the fact that a specific hazard on different jobs does exist and that this hazard is likely to vary by a rather large amount from one job to another.

This situation makes it necessary to rule out the element of job hazard in evaluating the concept of accident proneness. This was done in Table 46 by reassorting the data according to specific jobs. Table 46 is similar to Table 44 in general

interpretation. In it, however, the employees are divided according to the particular job they are on rather than the department in which they are working, as was done in Table 44. Thus, all openers, regardless of the department in which they are working, as we move from left to right across any

TABLE 45

VARIATION IN AVERAGE NUMBER OF HOSPITAL VISITS PER YEAR AMONG EMPLOYEES ON ELEVEN DIFFERENT JOBS

Job	*Average number of hospital visits per year*
Craneman	3.55
Opener	3.54
Reckoner	2.96
Machinist helper	2.77
Leader	2.75
Sheet inspector	2.54
Shear helper	2.40
Assorter	2.36
Potman	2.10
Foreman	1.16
Roll turner	.47

row, are exposed to approximately the same accident hazards. It may be seen in Table 46 that for nearly every job the average number of hospital visits for 1939 shows a progressive increase. The combined data for the eleven jobs considered are summarized in the bottom row of Table 46 and are portrayed graphically in Figure 127. We note nearly as definite a relationship between accidents of two successive years in Table 46 and Figure 127, when the element of job hazard has been essentially eliminated, as we previously noted in Table 44 and Figure 126 when the job-hazard element might conceivably have had a significant effect upon the results. From the data, it seems safe to conclude that accident proneness, or individual accident susceptibility, is significantly related to the type of accident that results in hospital visits for the plant studied.

It is one thing to determine, as the data above demonstrate, that an appreciable number of hospital visits are due to accident proneness or factors within the individual. It is

quite another thing, however, to identify the particular factors within the individual that determine accident proneness. In other words, once the concept of accident proneness has been established as an operating factor, the next question that arises is: What are the particular factors within an individual which tend to give him the particular degree of accident proneness that he is known to possess? It is obvious that two men could be equally prone to accident for entirely

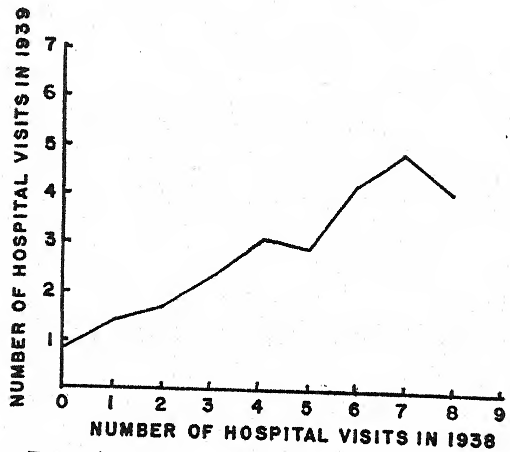


Fig. 127—Relation between number of hospital visits for two successive years among steel workers when hazard of the job is eliminated as a determining factor.

different reasons. Theoretically, at least, one man may be constantly getting into accidents because he is blind or nearly blind, while another man with perfect vision may find himself equally often involved in unsafe practices because of an unconscious desire on his part to escape certain responsibilities by experiencing occasional minor (or even major) injuries. The remedy in the first case might be as simple as obtaining a new pair of adequately fitted spectacles. The remedy in the second case, if it could be effected at all, might involve an extended series of psychiatric treatments. The point is that the specific factors underlying accident

TABLE 46

ACCIDENT PRONENESS FOR MEN ON JOBS WITH SIMILAR HAZARDS

Job	Average Number of Hospital Visits in 1939 for Men with 0, 1, 2, etc. Visits in 1938									
	Number of Visits in 1938									
	0	1	2	3	4	5	6	7	8	9
Craneman	.46	.92	.78	2.20	1.00	1.00	3.00	....	....	....
Opener	1.00	1.54	1.50	1.83	2.33	2.00	6.00	7.00	2.00	5.00
Reckoner	.65	1.42	1.40	2.00	3.20	3.00	....	6.00	8.00	....
Machinist helper	1.19	1.36	3.67	3.75	3.00	3.66	....	....	....	....
Leader	1.38	1.38	2.00	3.00	6.00	....	....	....	....	....
Sheet inspector	.50	1.85	.78	....	....	....	....	....	....	....
Shear helper	1.63	1.40	1.72	2.25	....	5.00	....	....	....	....
Assorter	1.48	1.10	1.75	2.60	2.80	2.00	3.41	1.50	2.00	3.00
Potman	.40	3.00	1.30	....	....	....	....	....	....	....
Foreman	.48	1.05	1.42	1.08	2.75	....	....	....	....	....
Roll turner	.48	.45	....	....	....	....	....	....	....	....
Average for all	.85	1.40	1.63	2.33	3.01	2.78	4.14	4.83	4.00	4.00

proneness must be determined before the plant management is in a position to make a really effective effort in the direction of reducing accident proneness of the high-accident employees.

External Factors Related to Accident Proneness

It is a recognized fact that many nonpsychological factors influence the accident rate and, to some extent, determine an employee's accident proneness. For example, it has been pointed out by Vernon[7] that the accident rate increases during the latter part of the working day. According to Vernon's results, this tendency is so marked that during a 12-hour working day women experienced $2\frac{1}{2}$ times as many accidents as during a 10-hour day. Although fatigue has often been considered the cause of this increase, the fact that the time of greatest accident rate as compared with hours worked is reversed on the night shift indicates that psychological rather than physiological factors are operating. Although Vernon's results attach somewhat more importance to the length of working day than do other investigations of this subject, it is quite commonly agreed that as the working day is lengthened, the accident rate increases in greater proportion than the increase in number of hours worked.

The job itself, as we have already noted, accounts for a part of accident proneness. Table 47 shows the hospital visits for ten departments of a sheet and tin mill. It will be noted in the last column of this table that the hospital visits per-man per year vary from a minimum of .55 to a maximum of 1.26. This difference clearly reflects the hazard of working in the different departments. Several investigators have found a good many other nonpersonal factors, such as heat, ventilation, and humidity, to be related to industrial accidents. Osborne and Vernon[8] report, for example, that

[7] H. M. Vernon, "An Experience of Munition Factories During the Great War," *Occupational Psychology*, XIV (1940), pp. 1–14.

[8] E. E. Osborne and H. M. Vernon, "The Influence of Temperature and Other Conditions on the Frequency of Industrial Accidents," *Industrial Fatigue Research Board*, Report No. 19 (1922).

whenever the temperature increases or decreases above or below a normal and comfortable level of around 70 degrees, the accident rate becomes correspondingly higher.

TABLE 47
DEPARTMENTAL DIFFERENCES IN HOSPITAL VISITS RATE

Department	*Number of Men*	*Number of Accidents for 1939 & 1940*	*Accident per Man per Year*
1	811	902	.55
2	573	1,144	1.26
3	480	723	.75
4	1,099	1,599	.73
5	336	555	.67
6	582	1,123	.96
7	624	1,238	.98
8	1,192	2,266	.94
9	373	529	.65
10	1,101	1,945	.88

It is also quite commonly accepted among safety men that an increase in the accident rate usually accompanies a step-up in production. Although this situation undoubtedly occurs in many individual plants when an attempt is made to step up production without adequate expansion of plant facilities, it is interesting to note that in a thorough study of the relationship between safety and production reported by the American Engineering Council[9] it was found, on a country-wide basis and over a period of years, that increased production in nearly every industry was accompanied by a decrease in the accident rate. This means that there is nothing inherently characteristic of a high-production rate that tends to make automatically for a high-accident rate. Rather it means that, as high production is achieved in the normal course of events by improved machinery and capacity for output, the accident rate will be likely to decrease along with the technological improvements and plant expansion. In any event, the effect upon safety of increased production, like the effect of temperature or job hazard, is essentially a physical rather than a

[9] American Engineering Council, *Safety and Production: An Engineering and Statistical Study of the Relationship Between Industrial Safety and Production* (Harpers, 1928).

psychological matter; and thus, important as this circumstance may be for the safety engineer who is interested in all angles of industrial safety, it is not primarily of psychological interest.

Psychological Factors Related to Accident Proneness

In contrast to the above-mentioned factors which partly determine accident proneness, certain factors related to accidents exist which are concerned primarily with the individual employee. Since in many cases these factors may be modified or even eliminated by methods that deal exclusively with the employee, they may properly be considered in the sphere of psychological factors that affect safety.

Vision

In spite of the fact that nearly every industry administers some form of vision test as a part of its employment procedure, vision surveys conducted among present employees in a number of industries reveal astounding differences in visual performance (see page 188). These differences are due in part to the changes in vision that occur with increasing age and in part to the fact that an injury to, or deterioration of, an employee's vision after he has been employed often may go unnoticed unless the injury takes place on company property during working hours. There are many reasons for conducting a periodic visual check-up of employees; one important reason is the relation that has been discovered between satisfactory vision and the frequency of accidents.

This problem has been investigated by comparing the accident experience of employees whose visual skills meet certain statistically determined standards[10] with the corresponding accident experience of employees whose vision does not meet these standards. The results of one study of

[10] The standards are determined by the method described in Chapter 7, page 212. When a safety visual profile is being set, the criterion against which the vision tests are correlated is low accident experience.

this sort are tabulated in Table 48, which shows the percentage of employees who sustained lost-time accidents over a two-year period in a group of several thousand steel mill employees unclassified as to job but segregated according to whether they passed or failed several vision tests. Those who failed had more accidents on the average than those who passed, except on the test of color vision. Such figures, however, need to be determined separately for each job in order to eliminate differences in factors of job hazard and job standards of vision that affect the total result.

TABLE 48
PERCENTAGE OF EMPLOYEES PASSING AND FAILING ON TESTS OF VISION WHO HAD SUSTAINED LOST-TIME ACCIDENTS

TEST	PERCENTAGE OF LOST-TIME ACCIDENTS		
	Meet Visual Standard	Do Not Meet Visual Standard	Ratio
Acuity, far	1.09	1.37	1.26
Depth perception	1.23	1.57	1.28
Color vision	1.21	1.14	.94*
Phoria	1.26	1.82	1.44

* Inversion in the data.

For this same unclassified group of employees the relationship between scores on a distance phoria test (in units of deviation from the mean) and serious accidents (home cases and lost-time cases) is graphed in Figure 66, which appears in Chapter 7. Employees showing a tendency to *converge* their eyes more than normal had experienced significantly more than the average number of serious accidents, while employees with a tendency to *diverge* their eyes more than normal had experienced fewer than the average number of serious accidents.

When a similar analysis was made of the relation between

near-phoria test results and accident experience, the opposite relationship was found. These results are shown graphically in Figure 67 in Chapter 7. In this instance the greater frequency of accidents was found among employees who showed a tendency toward divergence.

Another study reported by Stump[11] showed significant relationship between accident experience and several visual functions measured by the Ortho-Rater.[12] Other studies by Stump[13] have also shown significant relationships between vision and accidents. Wirt and Leedke,[14] studying paper-machine operators and tradesmen, found that employees with visual skills below a certain level experienced more serious accidents than employees with visual skills of greater adequacy.

The results of these several studies clearly reveal the relationship between vision and accident experience. When these statistics were first shown to safety men from one of the plants in question, a thorough study was made of the individual case records of the various lost-time accidents. Nowhere in the description of these accidents, which had been written at the time of the accident by persons familiar with the background of each mishap, could there be found reference to the fact that faulty vision had played a part in causing the accident. These individual case studies did not reveal a deficiency of vision of the employee responsible for the accident. Yet the statistics show that the visual factor is operating. On an actuarial basis, it can be predicted that, if all plant employees had been required to pass the critical vision tests, the lost-time accident rate among the employees

---

[11] N. F. Stump, "A Statistical Study of Visual Functions and Industrial Safety," *Journal of Applied Psychology*, XXIX (1945), pp. 467–470.

[12] The Ortho-Rater is described on pages 187 ff.

[13] N. F. Stump, "Visual Functions as Related to Accident Proneness," *Personnel*, XXI (1944), pp. 3–8; and "Industrial Safety and Visual Functions," *Journal of Psychology*, XX (1945), pp. 369–379.

[14] S. E. Wirt and H. H. Leedke, "Skillful Eyes Prevent Accidents," *Annual News Letter*, Industrial Nursing Section, National Safety Council, November, 1945.

would have been markedly decreased. The uncovering of the importance of this factor by statistical investigation, which had long escaped notice through ordinary case study of accidents, is a striking illustration of the value of the statistical approach in revealing hidden factors in any phase of human behavior.

No one should conclude from these statistics, however, that faulty vision is the *sole* cause of industrial accidents. Just as all kinds of factors may cause disease, so all kinds of factors may cause accidents. But the statistics do prove that faulty vision is *one* of the factors related to accidents. Certainly this factor is sufficiently important to deserve much more attention than it has received in many industrial plants in the past. Programs for such attention have been discussed in detail in Chapter 7.

Age and plant service

A number of the investigations of the relationship between accident rates and plant experience or job experience of the employee have been reported. As might be expected, these investigations do not reveal entirely consistent results from one plant to another or from one industry to another. In most of these studies it has been found that the accident rate is higher among the younger, more inexperienced employees. This is the conclusion reached by Hewes,[15] Gates,[16] Schmitt,[17] and Lipmann.[18] Previously unpublished data obtained in a sheet and tin mill also show a drop-off in the accident rate with the age of the employees, their years of service in the plant, and their years of service on their present jobs. These

---

[15] A. Hewes, "Study of Accident Records in a Textile Mill," *Journal of Industrial Hygiene*, III (1921), p. 6.

[16] D. S. Gates, "A Statistical Study of Accidents in Cotton Mills, Print Works and Worsted Mills of a Textile Company," *Journal of Industrial Hygiene*, II (1920), p. 8.

[17] E. Schmitt, "Unfällaffinität und Psychotechnik im Eisenbahndienst," *Industrielle Psychotechnik*, III (1926), pp. 144–153; 364–366.

[18] O. Lipmann, *Unfällursachen und Unfällbekampfung* (Berlin, 1925).

relationships are shown graphically in Figures 128, 129, and 130. Except for a low point of the curves for the very young or very inexperienced employees, there is a continuous decrease in hospital visits with increasing age or increasing plant or job experience.

Several possible explanations for the trends shown in Figures 128, 129, and 130 have been offered. It has been

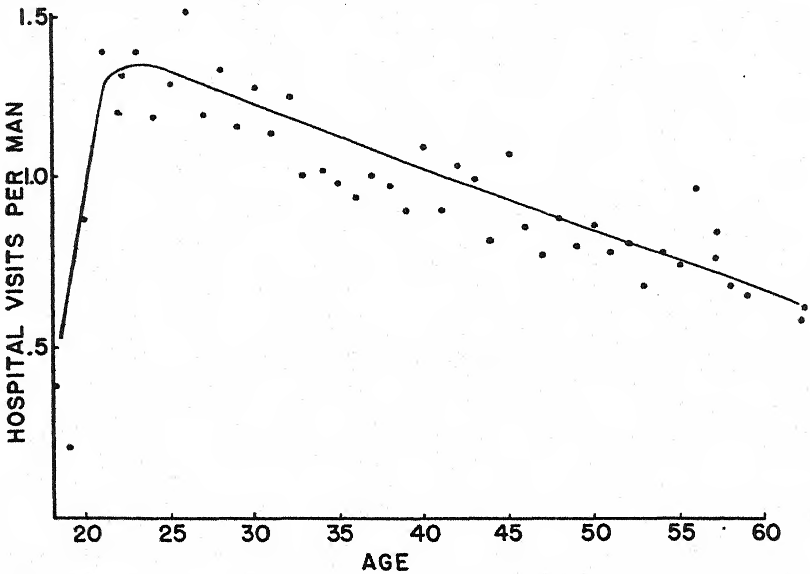


FIG. 128—Relation between age and hospital visits per man per year among 9,000 steel workers.

suggested that the younger employees are placed on the more hazardous jobs and that, as they become older and more "plant wise," they tend to be transferred to jobs of greater relative safety. It has also been suggested that the younger employee, who, in general, has relatively few family responsibilities, is less cautious and more likely to take chances than the older employee and therefore becomes involved in a greater proportion of industrial accidents. It has been further suggested that the younger employee is less acquainted with the possible dangers that exist in connection with the

different machines and is thus more likely to subject himself to situations in which an accident is likely to occur. Any or all of these factors (and possibly other factors which have not

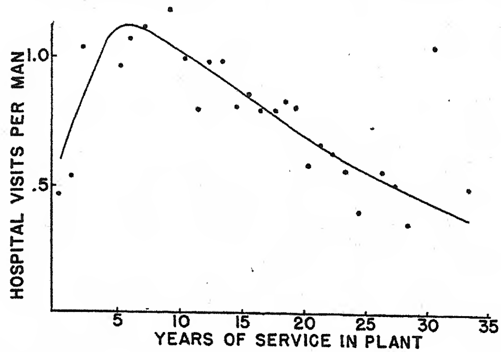


Fig. 129—Relation between years of service in the plant and hospital visits per man per year among 9,000 steel workers.

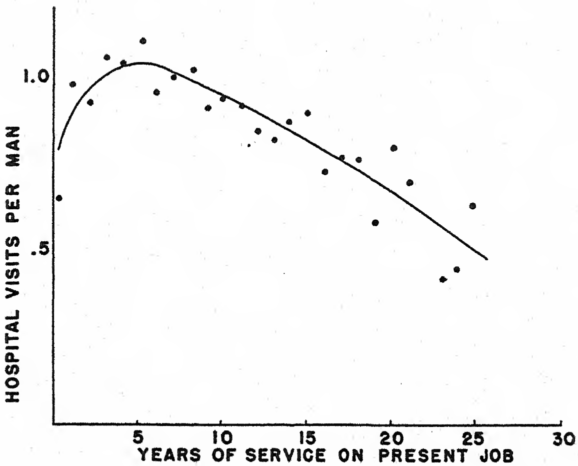


Fig. 130—Relation between years of service on present job and hospital visits per man per year among 9,000 steel workers.

been mentioned) may account for the results found. Whatever may be the cause or combination of causes, it is clear that in the plant studied the younger employee accounts for far more than his share of accidents. An adequate safety

program must therefore take into account the age and service differential in accident susceptibility. This can be done by safeguarding the younger employee while he is becoming experienced on the job, by organizing safety conferences with particular reference to these employees, and by making other efforts to instill in this high-accident group those habits of work and methods of thought which tend to reduce the hazard.

It should be emphasized, however, that the relationships described above are not necessarily true of industry in general and that in certain other investigations a reversed type of relationship has been found. Stevens,[19] for example, found in the plants which he studied that frequency of accidents increased with age and that the mature workman is disabled for a longer period than is the younger man when an accident does occur. Shrosbree[20] also reports statistics from one industry to show that workers of longer service are a greater hazard. This is due, according to Shrosbree, to their becoming accustomed to danger and hence less careful than the less experienced employee. Although the general rule seems to be a decrease in accidents with age, the last two studies cited indicate that this is not an invariable rule and that, occasionally, even the reverse principle may exist. The important thing to realize is not that a single rule necessarily applies to every industry, but that in any given industry the working conditions and other factors probably result in some sort of relationship. It is important for a safety director to know what relation exists in his plant and to direct his safety program specifically to that particular group of employees, young or old, who are working under the greatest accident hazard. Data of the type reported here are relatively easy

---

[19] A. F. Stevens, Jr., "Accidents of Older Workers: Relation of Age to Extent of Disability," *Personnel Journal*, VIII (1929), pp. 138–145.

[20] G. Shrosbree, "Relation of Accident Proneness to Length of Service," *Industrial Welfare* (1933), pp. 7–8.

to obtain in any industrial organization. As Howard[21] has emphasized, such data should definitely be in the hands of the safety director in formulating and administrating his safety program.

Emotional factors

Two emotional factors which have been found to be related to employee accidents are general emotional maturity and the emotional state at the time of the accident. An analysis of the causes of accident susceptibility among fifty motormen of the Cleveland Railway Company, published by the Metropolitan Life Insurance Company, is reproduced in Figure 131. In this analysis four single items, namely, faulty attitude, impulsiveness, nervousness and fear, and worry and depression, account together for 32 per cent of the accidents among the group studied. Since these four items are essentially emotional in nature, it may be inferred that emotional conditions account for a considerable proportion of accidents. It is well accepted psychologically that many individuals vary in their general emotional state between a "high" and a "low" condition. These conditions are often fairly regular in their occurrence and appear in a cycle which ranges from one extreme to another. This situation is readily recognized in its extreme form, as it results in the so-called manic-depressive psychosis. In a minor form, however, it characterizes many people who are in no sense of the word psychotic, but who do, nevertheless, experience definite "ups and downs" in their general emotional state. The relationship between these emotional cycles and the frequency of accidents has been shown in an investigation by Hersey,[22] who found that the average worker is emotionally low about 20 per cent of the time and that more than half of the four hundred minor

---

[21] R. R. Howard, "Importance of the Human Factor in Industrial Accidents," *Journal of American Insurance—Chicago*, II (1942), pp. 11–14.

[22] R. B. Hersey, "Emotional Factors in Accidents," *Personnel Journal*, XV (1936), pp. 59–65.

accidents studied occurred during these low periods. According to chance alone only 20 per cent of the accidents would have occurred during such periods; the differential may

PRIMARY CAUSES OF ACCIDENT-PRONENESS

Percentage Distribution Among Fifty Motormen

WOODHILL DIVISION, THE CLEVELAND RAILWAY COMPANY

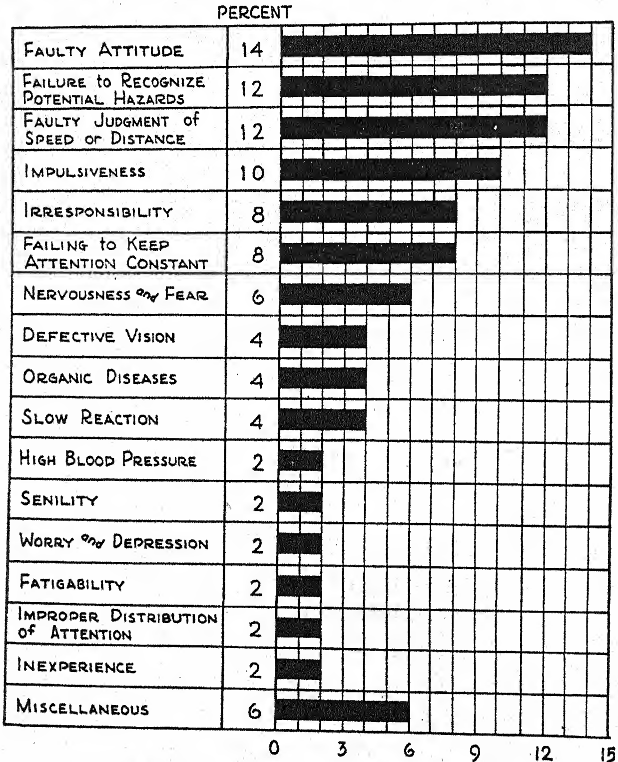


FIG. 131—Primary causes of accident proneness according to a survey among 50 motormen of the Cleveland Railway Company.

therefore be taken as indicating the effect that an emotional depression has upon an employee's safety.

The general importance of the employee's emotional

condition is further attested by another study by Hersey[23] in which it was reported that the production of industrial workers is around 8 per cent higher during periods when the men are elated, happy, hopeful, and co-operative than when they are suspicious, peevish, angry, disgusted, pessimistic, apprehensive, or worried. A favorable emotional condition thus not only is desirable from the standpoint of safety but also from the standpoint of plant productivity.

Supervisors and others dealing with human relations problems have long been aware of the fact that certain individuals vary from day to day in their agreeableness or ability to get along with others. Too often it has been the practice to handle such individuals either with a sharp reprimand or a general comment to the effect that so and so has a "grouch on" today. An employee who periodically gets a "grouch on" very likely is one who is afflicted with more than the average amount of cyclic fluctuation in his emotional state. To reprimand such an employee usually does about as much good as to reprimand a man with a broken arm for not doing his full share of work. Such an individual should be dealt with clinically to unearth the cause of his emotional condition. If this is impractical, he should be kept off of a job that is in any way hazardous for the duration of his unsatisfactory emotional state. It should be emphasized that an employee who suffers from extreme variations in mood is not doing so for the fun of it. He cannot help himself and he, himself, is unhappier as a result of his condition than are any of the people who are in contact with him.

Aside from the fairly common fluctuations in temperament that are quite certain to be found among a few at least of any large group of employees, there is the matter of general emotional maturity in responses which varies considerably from one person to another. It is entirely possible for an individual to be fifty years old by the calendar and yet be only

[23] R. B. Hersey, "Rates of Production and Emotional State," *Personnel Journal*, X (1932), pp 355–364.

five or six years old in terms of the level of his emotional behavior. A certain level of emotional development characterizes each state of the normal growth process. The baby cries and sleeps, the two-year-old kicks and screams, the eight-year-old fights, the adolescent sulks. The normal adult seldom indulges in any of these kinds of behavior; instead he has matured to a point where he recognizes and practices other, more social ways of gaining his end. Occasionally, however, we find a supposedly mature individual who has not grown up emotionally. He curses and fights like a child, sulks like an adolescent, and even occasionally cries like a baby. Such behavior obviously does not result in the degree of carefulness and responsibility that is necessary for safe practice in a modern industrial plant.

An indication of emotional characteristics of employees may often be obtained by means of personality tests such as those described in Chapter 6. Such tests, however, are of little value in measuring personality or emotional characteristics of an individual who does not want to be measured. Unfortunately, an unwillingness to be measured is often found in the very employees who are most in need of emotional development. The unwillingness itself is often a symptom of emotional maladjustment. Such cases require definite therapeutic measures of the type described by Anderson[24] and Wright.[25] Anderson describes several points to be followed in making a psychiatric study of an employee who has had emotional difficulties. Wright describes in some detail the counseling plan that has been in operation for a period of years in several plants of the Western Electric Company. The essential feature of this plan is that provision is made for individual study and counseling of problem employees by professionally trained counselors. Sometimes a counselor

---

[24] V. V. Anderson, "The Problem Employee, His Study and Treatment," *Personnel Journal*, VII (1928), pp. 203–225.

[25] H. A. Wright, "Personal Adjustment in Industry," *Occupations*, XVIII (1940), pp. 501–505.

will spend only a single period of a few minutes with an employee and will find that the problem of the employee has been satisfactorily solved. But often the counselor will find it necessary to devote a number of periods of one hour or even longer to a single employee before the trouble is really unearthed and the way made clear for a satisfactory solution. The work of these counselors might be described as a sort of applied psychoanalysis, although it differs in many ways from a really psychoanalytic treatment.

Some type of individual study and treatment seems definitely to be necessary in the case of those employees who are handicapped by severe emotional problems. Such treatment, while clearly of personal value to the employee, need not be looked upon by management as a sheerly paternalistic venture; on the contrary, it should be viewed primarily as a financial and business venture. The solution of the emotional problems of employees is a paying proposition in terms of reduced accidents and increased production.

Mental ability

Numerous investigations have shown that a certain minimum mental ability is required if an employee is to escape the hazards that are to be found in any industrial plant. Chambers[26] found that very few accident-prone individuals were above average in handwork, intelligence and learning ability, dependableness, and industry The accidents for the most part were contributed by individuals who were relatively low in the traits associated with mental ability. In another investigation, Henig[27] found a definite relation between accidents in the Essex County Vocational School at West Orange, New Jersey, and scores on the Army Alpha Mental

---

[26] E. G. Chambers, "A Preliminary Inquiry into the Part Played by Character and Temperament in Accident Causation," *Journal of Mental Science*, LXXXV (1939), pp. 115–118.

[27] M. S. Henig, "Intelligence and Safety," *Journal of Educational Research*, XVI (1927), pp. 81–87.

Test. The rapidly increasing use of some form of mental ability test as a standard part of employment procedure in many personnel offices is quite likely to be followed, as a secondary but very valuable result, by a reduction in accidents as well as an increase in general worker efficiency. The ever-increasing mechanization of modern industry is more and more increasing the importance of keeping away from the machine those individuals whose mental capacity renders them unable to recognize fully the dangers that are always present if unsafe practices are followed.

In apparent contradiction to the results described above, a few investigators, for example Farmer and Chambers[28] have found no correlation between degree of intelligence and accident liability. The disagreement is more apparent than real, however, for investigations of this latter type have attempted to correlate the degree of intelligence with accidents among employees who are above a certain minimum intelligence level. It seems clear that above a minimum critical mental ability level little, if any, relationship exists between further amounts of mental ability level and susceptibility to accidents. But it still remains quite necessary for an employee to possess this minimum amount of mental ability. Those who lack it are quite likely to be hazardous employees. It is primarily for the identification of this extremely low group that the use of mental ability tests is recommended as an important part of any accident-prevention program.

Ratio between perceptual and muscular speed

An interesting hypothesis which seems to account for at least a part of the cause of accident proneness has recently been advanced by Drake.[29] Drake administered to a number of accident-prone and safe individuals a series of psychological

[28] E Farmer and E G. Chambers, "A Psychological Study of Individual Differences in Accident Rates," *Industrial Fatigue Research Board*, Report No. 38 (1926)

[29] C. A. Drake, "Accident Proneness: a Hypothesis," *Character and Personality*, VIII (1940), pp. 335–341.

tests. The tests were roughly divided into two groups, namely, those dealing with perception and those dealing with muscular responses. The perceptual tests were primarily concerned with visual discrimination and the muscular tests with the speed of executing a number of routine manual activities. In the course of examining these test results it was observed that the accident-prone persons tended to have motor test scores which were relatively higher than their scores on the perception tests. It was also observed that the safer employees or accident-free individuals tended to have motor test scores which were lower, relatively, than their scores on the perception tests. Following this cue, a combination test score was obtained as follows:

$$\text{Test score} = \text{spiral test score} - \text{turning test score.}$$

In this formula the spiral test consisted of a hundred small aluminum spirals each of which had been punched with a small hole near one end. Fifty were punched "standard," that is, they contained the punched hole two and one half turns from the end. The remaining fifty were punched "off-standard," that is, the hole was punched at a distance from the end different from the two and one half turns of the standard punch. The person tested was required to separate the standard items from the others. Although an element of muscular dexterity was required in selecting these pieces, the test is primarily a measure of the individual's rapidity of perceptual discrimination in detecting the differences between the "standard" and "off-standard" pieces. The turning test referred to in the formula, on the other hand, involved the turning of ten pairs of machine screws into threaded holes in a vertical steel plate. This test, therefore, was largely one of speed of manual activity. The scores of both tests were first converted into comparable standard scores and then recomputed into a single final score according to the formula. It is clear from this formula that a positive final score results when the perceptual test score is higher than the muscular test score,

and a negative final score results when the opposite situation with reference to the two initial tests prevails.

After the composite test score obtained in this manner had been obtained for each of thirty-eight employees, the accident index of each employee was computed according to the following formula:

$$\text{A.I.} = \frac{\text{Number of Accidents} \times \text{Severity}}{\text{Length of Service in Months}}$$

The next step was to plot the accident indices so obtained for the 38 employees against the composite test score. This

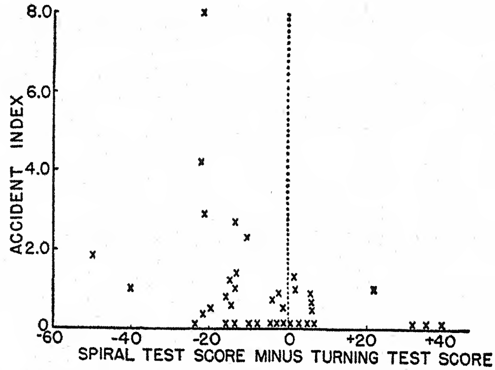


FIG. 132—Relation between accident proneness and test scores dealing with perceptual and muscular speed.

has been done in Figure 132. The general trend of the relationship is clear. Though it is not a straight line or linear relation, individuals with negative composite scores show a definite tendency to be among those with the high accident indices, whereas those with the positive composite scores are relatively free from accidents. Drake's statement of the principle involved, as illustrated in Figure 132, is that "Individuals whose level of muscular action is above their level of perception are prone to more frequent and more severe accidents than those individuals whose muscular actions are below their perceptual level. In other words, the person who reacts quicker than he can perceive is more likely

to have accidents than is the person who can perceive quicker than he can react."

Other psychological factors

Other attempts have been made to identify the accident-prone employee by means of tests; notable among these is the work of Farmer and Chambers.[30] These investigators administered to a number of industrial employees a group of muscular tests (dotting tests, reaction time tests, and pursuit meter tests), a group of what were called temperamental instability tests (including muscular balance, rate of tremor, and psychogalvanic reflex), and tests of the higher thought processes (which included intelligence and a number-setting test).

Follow-up studies were then made to determine whether employees who were above average in these tests were, in general, more free of accident records than employees whose test scores were below average. For six separate groups of employees, it was found, those whose performance on the dotting tests was better than average had a more favorable safety record than the poorer testing employees. Apparently the muscular skill required to obtain a favorable score on the dotting test is, in some way, of value to an employee in helping him to avoid accidents in the plant.

No one would contend on the basis of the preceding studies that any single one of the factors discussed is the only one related to industrial accidents. But the analyses indicate that, in the studies cited, each factor is related in some measure to plant accidents and that, together, these identified factors are associated with a significant proportion of such accidents. It seems clear, then, that individual safety can never be fully attained by improvement of working conditions and safeguarding the machines. No matter how safe a machine is, the accident-prone employee may find himself injured by it. An employee whose vision is unsatisfactory

---

[30] Farmer and Chambers, *op. cit.*

can severely injure himself on a machine that is perfectly safe for an employee with normal vision; and an employee with a temporary or continuing maladjustment may be injured even when the machines with which he is in contact are provided with every known safeguard. The individual must be considered, for the causes of many accidents are to be found only by exploration of his make-up. The adequate handling of the accident-prone employee is just as important a part of the safety engineer's job as is the installation of machinery safeguards.

How to Reduce Accident Proneness

Several ways to reduce the proportion of accident-prone employees among newly hired employees and to reduce the degree of accident proneness among present employees may be seen to follow logically from the foregoing statement of factors that are related to accident proneness.

Psychological tests at the time of employment

When it has been found that accident-prone employees and safe workers make significantly different scores on a psychological test, that test may safely be used at the time of employment to detect and reject those applicants who are in the unsafe or accident-prone category. Rigid adherence to such a policy of hiring in the employment office is as beneficial to the rejected employees as to the management. When employees are hired without reference to those psychological and personal factors that are related to accident proneness, the end result is the placement of employees in such a manner that an unnecessary number of accidents occurs. For example, it has been demonstrated (see page 219) that employees who fail certain vision tests experience more industrial accidents than employees who pass those tests. Therefore, the hiring of employees without regard to their scores on vision tests is almost certain to be followed by a poorer safety record than could easily be obtained if certain critical scores on these tests

were made one of the requirements for employment. Many, perhaps most, industries give some sort of vision test as a part of employment procedure. But often this test consists only of the Snellen Chart at twenty feet, and this test measures no more than visual acuity at that distance. Among employees passing this test are many who experience accidents because of visual trouble (such as lateral or vertical phorias) that is not detected by the Snellen test. A comprehensive visual performance test is necessary as a part of employment procedure if the employees who are accident prone for visual reasons are to be identified. Many, if not most, failures on visual performance tests can be remedied by proper professional attention. But without the stimulus of vision tests at the time of employment, many applicants would be quite willing to go to work with vision entirely inadequate for safe conduct within the plant.

Other tests dealing with such factors as mental ability and muscular speed may also be used to advantage, particularly in selecting employees for the more hazardous jobs. The several investigations summarized on the relation between mental ability and accident experience all point to the general conclusion that a minimum score on an adequate mental ability test should be required of employees who are to be exposed to the normal hazards of a modern industrial plant. This does not mean that a particularly high critical score need, or should, be required for many jobs; but it does mean that a reasonable minimum score should be required and that an applicant who is so low in mental ability as to fall below that score should not be employed.

Safeguarding and training inexperienced employees

The fact that young or inexperienced employees ordinarily experience more than their share of industrial accidents suggests the need for special care in building up a "safety consciousness" in the members of this group. This may be done in several ways. New employees often may be placed

for a time on jobs that are not hazardous but that are close enough to those jobs where great care must be exercised that the new employee will have an opportunity to learn by observation something about the machinery with which he later will be in contact. This procedure enables the new employee to have a "breathing spell" during which he can adjust himself to the plant and overcome whatever feelings of uncertainty and anxiety he may have immediately after going to work in a strange (and, to him, often awesome) industrial plant.

Another method of helping the new employee is to assign him to periodic safety conferences. Conferences of this type are becoming more and more widely used as a tool of management, particularly on a supervisory level, and more is said about the general nature and purposes of such employee conferences on page 275. These conferences are of real value as one means of assisting the new employee in his task of becoming familiar with the hazards of his job. Every safety engineer has available a list of unsafe practices that apply to the several departments and operations of his plant. These unsafe practices are usually fairly well known to experienced employees who have seen men hurt when the safety rules were not followed, and who, therefore, are likely to avoid the unsafe practices with reasonable care. But the new employees, who often do not know what the unsafe practices are, are likely to endanger themselves not because they are essentially careless or foolhardy, but simply because they do not know *how* to be safe.

Posters and placards may help the new employee to some extent, *if* the posters tell him specifically *what* to do and what *not* to do, and *if* he reads the posters and remembers what he reads. But too often posters simply show an accident and, in large type, say "Be Careful." Since only the maladjusted individual is likely to hurt himself voluntarily, the degree of carefulness of the ordinary employee is determined by (1) what he knows of safe and unsafe practices, and (2) how

important he thinks it is to follow his knowledge on the subject. In other words, some employees follow unsafe practices because they actually do not know any better. The need in such cases is for specific information on what to do and what not to do. A systematic and comprehensive series of conferences covering this specific information is a much more effective means of informing the employees than is the chance and casual learning obtained from posters or from observations that employees may (or may not) make. Once the information has been given to the employees, still further education is necessary to make sure that they *believe* as well as *know* that certain practices are unsafe.

Personal protection

The use of recommended safety measures, such as safety goggles and safety shoes, might be considered primarily as physical rather than psychological equipment, if it were not for the fact that the use an employee makes of such safeguards is so often determined by his own personal beliefs, attitudes, and, sometimes, prejudices. Many men believe that safety goggles are unnecessary even on jobs that are demonstrably dangerous because "they have never had an accident." This is one instance in which neither management nor employees can safely follow the adage that experience is the best teacher. Those who wait for experience to teach them that goggles should be worn to protect their eyes may have no eyes to protect after they have learned the lesson. The adequate utilization of safety measures requires not only that such measures be made available to the employees but also that the employees be educated in the use of these measures and *convinced* of their value. It is just as much a part of management's job to insure the success of the educational part of the program as to supply the safety measures.

Job analysis and accident records often reveal that employees on certain jobs are particularly liable to visual injury. However, until such studies have been made,

management is not always aware of the particular departments or jobs in which a visual hazard is most likely to exist. Table 49 shows a breakdown by departments of the relative frequency of seven different types of minor injuries experienced during two years by approximately 9000 employees in a steel mill. It was noted that in one of the ten departments, Department 7, visual injuries were more frequently experienced than any other type of accident. In eight of the ten departments visual injuries were second in frequency, and in one department the visual injuries were third in order of frequency. For the plant as a whole, eye injuries were second in importance, being exceeded in frequency only by cuts and bruises. Many of the eye injuries reported were of course of a minor nature, perhaps involving little more than careful removal of a foreign body from the eye. However, even injuries of this nature would be better prevented, for they involve not only loss of time for the employee and the company but also the possibility of serious complications requiring compensation and rehabilitation.

Influenced by such evidence as that presented in Table 49, many industries have adopted a mandatory safety-goggle program. Such a program requires the wearing of safety goggles by everyone in the plant at all times and regardless of his job. The experience of organizations that have adopted such a program, such as the Pullman Company,[31] has been highly favorable. Indeed, these companies report a marked financial saving that is directly traceable to the reduction in eye accidents effected by the safety-goggle program.

Immediate attention to injury

Management finds it a paying proposition to be sure that every injury, no matter how slight, receives prompt and adequate attention. It is not enough to have hospital or

[31] H. Guilbert, "No One Enters This Plant Without Goggles," *Factory Management and Maintenance*, XCVI (1938), pp. 83–85.

TABLE 49

Departmental Differences in Type of Accidents

DATA BASED ON MINOR ACCIDENTS FOR ONE YEAR IN A STEEL MILL

Type of Injury	Department									
	1	2	3	4	5	6	7	8	9	10
Eye Injuries	14.4%	13.8%	14.1%	11.3%	25.5%	19.4%	38.8%	12.7%	20.4%	24.8%
Strains	3.5	3.3	2.9	2.6	3.4	2.6	3.7	3.2	3.5	2.7
Fractures	1.5	0.4	0.3	0.1	1.1	0.5	0.8	0.8	1.2	0.1
Heat	1.8	0.7	15.7	0.6	1.2	0.3	5.3	0.9	0.1	0.1
Burns	4.9	3.5	9.7	4.1	5.2	3.7	9.1	3.7	5.2	6.5
Cuts & Bruises	64.3	71.3	50.0	73.6	54.1	66.5	37.1	70.7	60.0	55.7
Miscellaneous	9.6	7.0	7.6	7.4	9.7	7.1	5.5	8.0	9.6	10.1
Total Number of Hospital Visits 1938–1939 (Minor Injuries)	1217	1390	861	1852	652	1416	1433	3000	657	2269

first-aid service available for an employee to make use of if he wishes to do so. Means must be found to insure his making use of such facilities. Very often factors that have been totally overlooked by management effectively prevent many employees from taking advantage of available first-aid facilities. In one mill, for example, two machine shops contained identical machinery, employed approximately the same number of men, and turned out the same kind of work. An analysis of the accident records at the end of a year showed that the employees in one of these shops averaged 1.73 hospital visits per man for the year, whereas the employees in the other shop averaged only 1.10 hospital visits per man for the same period. So great a difference in the apparent hazards of the two shops was entirely unexplainable in terms of the men, the work, or the shops themselves. After thought was given to the problem, it was remembered that the hospital in which first aid and treatment were administered was next door to one of the machine shops but was about five hundred yards from the other. The men in the shop close to the hospital took advantage of the service in nearly 60 per cent more cases of minor injury than did employees in the more distant shop. The solution to this problem is in educating the employees in the latter shop to walk five hundred yards to receive treatment when treatment is needed, or locating another emergency first-aid station closer to these men. Management should not and cannot take the point of view that the service is there and that employees who do not take advantage of it are on their own responsibility. Management must see that the service is utilized if a satisfactory safety record is to be obtained.

14

Attitudes and Morale

In an army, a school system, or an industrial plant, there is no substitute for morale. The main difference between men and machines is that the productivity of a man is determined very largely by the way he feels about his job and the other employees with whom he works, and by his attitude toward the company that employs him. There is an old saying that loyalty has no price. One can hire "hands" but the "hands" are of little value if the man who owns them feels that he is not being treated fairly by the management.

The very use of the term *hands* in referring to employees (commonly encountered a few years ago and still prevalent in rural areas) reflects a mistaken viewpoint by management, as a man's hands alone are never hired. The whole man is always hired, and the whole man brings to work a good many things besides his hands. He brings the effects of too many or too few spankings as a child; of whether he won or lost in last night's card game; of whether his wife sent him to work with a scolding or a kiss; of whether or not the company "docked" his wages when he was home last week with a sick child. Such things as these are of vital importance in determining an employee's real value to a company. And such things as these, considered together for the whole working force, determine the morale of an industrial plant.

Morale cannot be legislated or induced by logical argument; neither can it be bought for a price. The employees of many industries with wage scales above average for the locality are constantly bickering because they feel that "they are

not being treated fairly by the company." Management has often made the mistake of assuming that high wages and short hours are the only things men want from their jobs although studies such as the one mentioned on page 458, as well as other surveys summarized later in this chapter, show that many non-monetary factors are considered of great importance by both organized and unorganized working forces. When a company assumes that monetary factors and hours of work are the sole or basic reason for employee unrest, an attempt has often been made to remedy the situation by raising wages or shortening the hours of work, or both. And then management has been amazed by a continuation of the discontent and often has been quite unable to understand why the men do not return happily to work.

Nor has the lack of understanding what men really want from their jobs been limited to management. Many employees are themselves unaware of what is needed to make them satisfied with their jobs. This does not mean that the employees are unintelligent or that they have any special disability to analyze their true motives. It simply illustrates a well-established principle of normal human behavior, namely, that it is difficult for anyone to identify in clear and unmistakable terms the forces that underlie his feelings and his actions. Everyone has observed the employee who, under one supervisor or in one working group, continually sulks on the job because "the company doesn't pay him what he is worth," and, under another supervisor or in another working group, works in a contented and industrious manner for exactly the same wage. It is not unusual to watch a man's attitude shift from one of chronic discontent to one of complete satisfaction following a shift from a job of low status to one higher in this intangible quality, even though the change in jobs involved no appreciable increase in wages.

Every impartial study of what industrial employees want from their jobs has shown that many things besides high wages and short hours are desired. Such studies have shown,

also, that some employees consider these other things even more important than wages and hours.

What Men Want from Their Jobs

Typical of studies on this subject is one by Houser,[1] who found that the nonselling employees (including unskilled

TABLE 50

Relative Importance of 28 Factors to the Nonselling Employees of a Large Merchandising Organization

	Rank
Receiving help necessary to get results expected by management	1
Being encouraged to offer suggestions and try out better methods	2
Being able to find out whether work is improving	3
Reasonable certainty of being able to get fair hearing and square deal in case of grievance	4
Certainty of promotions going to best qualified employees	5
Encouragement to seek advice in case of real problems	6
Being given information about important plans and results which concern the individual's work	7
Being given reasons for changes which are ordered in work	8
Not being actually hampered in work by superior	9
Getting contradictory or conflicting orders	10
Being given to understand completely the results which are expected in a job	11
Pay—*Assurance of increases when deserved*	12
Being invited to offer suggestions when new plans are being considered	13
Feeling that superior understands all about the difficulty of the individual's job	14
Being given to understand completely the general methods which the superior wants followed	15
Complete definition of duties	16
Not being responsible to too many superiors	17
Knowledge of other jobs in the organization which the individual feels capable of handling and would prefer	18
Knowledge of other jobs preferred, even at same pay	19
Red tape in the organization, preventing best work	20
Pay—*compared to that of other jobs of equal importance in the organization*	21
Pay—*compared to that of similar work in other organizations*	22
Treatment when being employed	23
Knowledge of lines of promotion	24
Value of Mutual Benefit Association	25
Being permitted to make important decisions in work	26
Regularity of amount of work	27
Service of Medical Department	28

---

[1] J. D. Houser, *What People Want from Business* (McGraw-Hill Book Company, 1938), p. 29.

labor) of a large merchandising organization consider the factors shown in Table 50 to be the things they wanted from the company. The order in which these factors are listed in Table 50 shows the relative importance attached to the factors by the employees. It is interesting to note that the employees of this company ranked the items concerned with pay twenty-first and twenty-second in importance.

TABLE 51

PERCENTAGE OF UNION AND NONUNION EMPLOYEES WHO CONSIDERED THE LISTED ITEMS AS "MOST IMPORTANT" AMONG FACTORS RELATED TO THEIR JOBS

	Per Cent of Union Employees Checking Item as "Most Important"	*Rank of Item*	*Per Cent of Non-Union Employees Checking Item as "Most Important"*	*Rank of Item*
1 Employee stock subscription...	5	11.5	2	13.5
2 Voice or share in management..	13	9.5	6	11.0
3 Fair adjustment of grievances..	80	1.0	24	7.0
4 Chance of promotion..........	28	6.0	47	3.0
5 Steady employment...........	65	2.0	93	1.0
6 Medical and dental service.....	0	13.5	6	11.0
7 Safety.......................	57	3.0	21	9.0
8 Amount of pay...............	49	4.5	51	2.0
9 Working conditions...........	49	4.5	45	4.0
10 Hours of work................	13	9.5	23	8.0
11 Type of man in charge........	18	7.5	38	5.0
12 Methods of pay...............	0	13.5	2	13.5
13 Insurance systems and pensions	18	7.5	36	6.0
14 Chance to show intiative......	5	11.5	6	11.0

Though studies of other employees in other plants do not consistently show the financial incentive to be so far down the list as it appears in Houser's study, other studies do almost universally show that non-monetary factors are of significant importance. Hersey,[2] in a study of union and nonunion employees, found that fourteen factors were considered "most important" by the percentages of employees tabulated in Table 51.

Hersey did not find the financial aspect of the job to be

[2] R. B. Hersey, "Psychology of Workers," *Personnel Journal*, XIV (1936), pp. 291–296.

quite so unimportant, in comparison with other factors, as did Houser in his study. Yet Hersey did find two things of importance. First, he found that amount of pay was not considered most important by most employees in either the union or the nonunion group. "Steady employment" was checked by more employees as a most important factor in both groups and, in the union group, the factors of "fair adjustment of grievances" and "safety" also outranked "amount of pay." Second, Hersey found that still other factors, though admittedly of somewhat lesser importance than amount of pay to the majority of employees, are considered of first importance by a sufficiently large percentage of employees to justify definite attention by management.

TABLE 52

RANKING OF TEN ITEMS IN ORDER OF IMPORTANCE BY 325 FACTORY WORKERS

Rank	*Factor*
1	Steady work
2	Comfortable working conditions
3	Good working companions
4	Good boss
5	Opportunity for advancement
6	High pay
7	Opportunity to use your ideas
8	Opportunity to learn a job
9	Good hours
10	Easy work

Another investigation dealing with the factors that appeal to factory workers has been reported by Wyatt, Langdon and Stock.[3] In this study, ten factors were ranked by the 325 employees in the order of importance shown in Table 52.

In this report, as in those mentioned before, many factors besides those of a monetary nature are considered of real importance by the industrial employees studied.

Hoppock[4] has published the results of an interesting investigation into the problem of job satisfaction and the conditions that favor it. In one part of the study, eighty

[3] S. Wyatt, J. N. Langdon, and F. G. L. Stock, "Fatigue and Boredom in Repetitive Work," *Industrial Health Research Board*, Report 77 (1937).

[4] R. Hoppock, *Job Satisfaction* (Harper & Brothers Company, 1935).

persons representing a wide range in age, intelligence, occupations, and earnings were interviewed. From these results Hoppock states that no conclusions are drawn except "that what happened in these cases can happen." However, the interview case studies corroborate the general conclusion reached above, that is, that job satisfaction is related to a good many things besides financial return. Some of these factors, as summarized by Hoppock, are "relative status of the individual within the social and economic group with which he identifies himself, relations with superiors and associates on the job, nature of the work, earnings, hours of work, opportunities for advancement, variety, freedom from close supervision, visible results, the satisfaction of doing good work, opportunities for service to others, environment, freedom to live where one chooses, responsibility, vacations, excitement, opportunity for self-expression, competition, religion, opportunity for or necessity of traveling, fatigue, appreciation or criticism, security, and ability to adjust oneself to unpleasant circumstances."

Further evidence of the importance of non-monetary factors in influencing employee morale comes from a study reported by Roberts,[5] who made an audit of employee morale of 2,500 employees. His findings show the importance to morale of such factors as promotion procedures, favoritism, and the effective informing of employees of company policies. A study reported by Blum[6] also showed that certain non-monetary factors—job security and advancement—ranked above pay in their influence upon employee morale.

These studies should not be interpreted by management as meaning that the average employee has no interest in his pay check or that he wants to be "fathered" or "mothered"

---

[5] E. B. Roberts, "Tests to Determine Objectively the Effectiveness of an Industrial Relations Program," *Office Management Series, American Management Association*, No. 84 (1938), pp. 32–37.

[6] M. L. Blum and J. J. Russ, "A Study of Employee Attitude Toward Various Incentives," *Personnel*, XIX (1942), pp. 438–444.

or in any other way made the beneficiary of a paternalistic system. The average employee wants to take care of himself. He does not want the company to pay his medical or dental bills (see item 28 in Table 50 and item 6 in Table 51), nor is he particularly interested in being given a voice in the management of the plant (see item 26 in Table 50 and item 2 in Table 51). He wants the opportunity to do a job for a reasonable wage, to have an impartial hearing if he thinks he has been treated unfairly, and to be reasonably sure of holding his job as long as his work is satisfactory. In one investigation,[7] the last-mentioned point, job security, was found to be more important than any other factor.

Added to these factors should be an opportunity for the employee to work with a group that "accepts" him and of which he feels himself to be an integral part. A series of investigations conducted by Mayo[8] and his collaborators in a variety of plants has shown the influence of social and group factors upon the worker's production, quality, absenteeism, and related characteristics. These studies show, among other things, that many of the factors (such as illumination) previously considered to be specific *determiners* of employee productivity operate in a much less specific and tangible fashion than has previously been supposed. For example, in a series of studies[9] of a small group of employees at the Hawthorne plant of the Western Electric Company, an effort was made to determine the effect upon production of varying the illumination. When the lighting was increased, production increased; when the lighting was left the same, production also increased; and when the lighting was *decreased*, production still increased. In this particular experiment, whatever was done had a favorable effect on employee job

[7] D. McGregor, "The Attitudes of Workers Toward Layoff Policy," *Journal of Abnormal and Social Psychology*, XXXIV (1939), pp. 179–199.

[8] Several of these studies have been summarized by Elton Mayo, *The Social Problems of an Industrial Civilization* (Andover Press, 1945).

[9] Reported by Whitehead, T. N., *The Industrial Worker* (Harvard University Press, 1938).

performance. The only reasonable explanation seemed to be that the employees in the experimental group considered themselves a rather special group of employees, and that they were responding in a favorable manner to whatever changes management made in their working conditions. There is little reason to believe that such factors as a new job evaluation plan, a new incentive plan, a new labor contract, or the careful selection and training of new employees would have entirely taken the place of this *group feeling* in improving job performance. But factors affecting employee *morale* might quickly undermine this group feeling; these factors, therefore, must be considered by management if employee morale is to be kept on the high level that efficient plant operation requires.

The Measurement of Morale

Morale cannot be weighed on a scale, like a pound of butter; it cannot be measured with a rule, like a strip of carpet; it cannot be gauged with a thermometer, like the temperature of a room. Yet, though it is intangible, it can be measured. Thorndike long ago pointed out that whatever exists, exists in some amount, and whatever exists in some amount can be measured. Certainly morale does exist. It is a characteristic of a group of employees that makes for success or failure of the plant as a business enterprise. Until recently, management attempted to keep in touch with the morale of the working force only through such factors as chance remarks by employees, appearance and behavior of the men at work, and occasional reports by supervisors. These methods were none too satisfactory. It is one characteristic of an employee whose morale is poor that he is likely to keep his attitude to himself, particularly if he feels that he personally is likely to be identified with a condition of dissatisfaction. Therefore, if morale is to be gauged or measured, means must be found to encourage employees to express their honest feelings and reactions. One satisfactory method of doing this is by means

of an anonymous employee-attitude survey. Such a survey makes use of a questionnaire adapted to the needs of the investigation. The employees are asked to check the questionnaire and drop it in a ballot box or mail it to a neutral center, frequently a university or management consulting organization. No one knows how any particular employee checked the questionnaire, but from the average results, a comprehensive picture of general employee reactions is often obtained. Attitude surveys of this sort may make use of a single "attitude toward the company" scale, or of a specially prepared questionnaire designed to determine employee opinion on a number of specific topics of interest to the management.

Attitude scales

A typical example of an attitude scale that may be used to determine the general attitude of employees toward their

TABLE 53

STATEMENTS USED IN BERGEN'S SCALE FOR MEASUREMENT OF ATTITUDE OF EMPLOYEES TOWARD THEIR COMPANY

	Scale Value
I am made to feel that I am really a part of this organization	9.72
I can feel reasonably sure of holding my job as long as I do good work	8.33
I can usually find out how I stand with my boss	7.00
On the whole, the company treats us about as well as we deserve	6.60
I think training in better ways of doing the job should be given to all employees of the company	4.72
I have never understood just what the company personnel policy is	4.06
In my job I don't get any chance to use my experience	3.18
I can never find out how I stand with my boss	2.77
A large number of the employees would leave here if they could get as good jobs elsewhere	1.67
I think the company's policy is to pay employees just as little as it can get away with	.80

company is shown in Table 53. This scale is taken from an article by Bergen.[10] The theory of such scales and detailed instructions for their construction have been described by

[10] H. B. Bergen, "Finding Out What Employees Are Thinking," *The Conference Board Management Record*, April, 1939.

Thurstone and Chave.[11] Such a scale requires that numerical values be found for each of a series of statements that express a thought (favorable, unfavorable, or indifferent) about the subject of the attitude survey.

The numerical values assigned to the statements in Table 53 are not determined arbitrarily. They are obtained from a careful series of experiments conducted during the construction of the scale. In the scale construction, the first step is to write out a large number of statements, perhaps a hundred or more, each of which expresses a viewpoint of some kind toward the company. An effort should be made to have these statements express all possible attitudes from extremely favorable to extremely unfavorable. Each of these statements is typed on a separate slip of paper and a judge is asked to place each statement in one of ten piles, ranging from statements showing the least favorable viewpoints, in pile one, to statements expressing the most favorable viewpoints, in pile ten. When the judge has allocated all statements, a tabulation is made of where he placed each one, and the statements are shuffled and given to another judge. This procedure is repeated until the statements have been separately allocated to the various piles by approximately one hundred judges. It should be emphasized that the hundred judges used as described above are assisting in the construction of the scale. They are not having their own attitude measured. To measure attitudes with the scale is not possible until the scale has been constructed, and the allocation of statements to the several piles is a part of the process of constructing the scale.

The purpose of the allocation is to determine the scale values of the various statements. If all judges tend to place a statement in piles toward the favorable end of the continuum, we may safely conclude that that statement expresses a favorable attitude toward the company. If a statement is

[11] L. L. Thurstone and E. J. Chave, *The Measurement of Attitude* (University of Chicago Press, 1929).

generally placed by the judges in piles toward the unfavorable end of the series, we may likewise conclude that an unfavorable attitude is expressed by that particular statement. The piles are numbered from the unfavorable to the favorable end of the continuum. The number of times each statement is placed in each pile is determined, and a computation made to determine the average location of the statement by the judges. From this calculation, the scale value of the statement is determined. By starting with many more statements than need be retained for the final attitude scale, it is usually possible to pick from ten to fifteen statements that are spread over the entire attitude range. The statements so selected, together with their scale values as shown in Table 53, comprise the final material for the attitude scale.

One might think that the attitudes or feelings of the judges who are used in the construction of the scale would have an effect on the scale values obtained. In other words, it might be felt that one set of scale values for a series of statements might be obtained if the judges were, in general, *favorable* toward the company, while a different set of values might be obtained if the judges in general were unfavorable or indifferent toward the company. This possibility has been subjected to experimental test, and it has been found that the attitudes of the judges do not significantly affect the scale values obtained from them.[12] This fact increases the possibilities of use for such scales in industry because it often happens that the persons who are most conveniently available for use as judges in constructing a scale may be more favorably disposed toward the company than certain groups of employees with whom the scale is to be used after it is constructed.

In the practical administration of an attitude scale, statements are printed on a sheet in random order, without the scale values appearing in Table 53. Each employee is given one of these sheets and is requested to check all statements

---

[12] E. D. Hinckley, "The Influence of Individual Opinion on the Construction of an Attitude Scale," *Journal of Social Psychology*, III (1932), pp. 283–296.

that he agrees with or believes to be true. The sheets are then turned in without being signed. The attitude of an employee toward the company is defined as the average scale value of the statements he has checked. For example, an employee checking statements one, three, and five of those shown in Table 53 would have an attitude of

$$\frac{9.72 + 7.00 + 4.72}{3} \quad \text{or} \quad 7.15.$$

On a scale of ten (ten being the most favorable end and zero the least favorable end) an attitude of 7.15 would be one somewhat toward the favorable end of the scale. On the other hand, an employee checking statements seven, eight, and ten would have an attitude represented by

$$\frac{3.18 + 2.77 + .80}{3} \quad \text{or} \quad 2.25$$

This would be a much less favorable attitude toward the company than the one described above.

The attitude scale reproduced in Table 53 is by no means the only set of statements that might be constructed to fulfill the requirements of a suitable scale. Uhrbrock,[13] in a comprehensive study of the attitudes of 3,934 factory workers, 96 clerical workers, and 400 foremen, used a different scale, but one constructed as explained above to measure the general attitude of employees toward the company. Some typical statements, with their scale values from Uhrbrock's scale, are reproduced in Table 54.

The fact that scales such as those illustrated in the Tables 53 and 54 may be used independently to measure the attitude of employees toward their company illustrates several points concerning the use of such scales. It will be noticed that the statements comprising the two scales are not the same, yet both scales measure essentially the same thing. Thus, the

[13] R. S. Uhrbrock, "Attitudes of 4430 Employees," *Journal of Social Psychology*, V (1934), pp. 365–377.

use of such scales does not require including any specific statements which, from an industrial relations viewpoint, might seem undesirable. A scale can be "tailor made" for a given plant and deliberately kept free of statements that may seem to "pack dynamite" without sacrificing the validity

TABLE 54
STATEMENTS USED IN UHRBROCK'S SCALE FOR MEASURING ATTITUDE OF EMPLOYEES TOWARD THEIR COMPANY

Statement	*Scale Value*
I think this company treats its employees better than any other company does	10.4
If I had to do it over again I'd still work for this company	9.5
They don't play favorites in this company	9.3
A man can get ahead in this company if he tries	8.9
I have as much confidence in the company physician as I do in my own doctor	8.7
The company is sincere in wanting to know what its employees think about it	8.5
A wage incentive plan offers a just reward for the faster worker	7.9
On the whole the company treats us about as well as we deserve	7.4
I think a man should go to the hospital for even a scratch, as it may stop blood poisoning	6.3
I believe accidents will happen no matter what you do about them	5.4
The workers put as much over on the company as the company puts over on them	5.1
The company does too much welfare work	4.4
Soldiering on the job is increasing	4.1
I do not think applicants for employment are treated courteously	3.6
I believe many good suggestions are killed by the bosses	3.2
My boss gives all the breaks to his lodge and church friends	2.9
I think the company goes outside to fill good jobs instead of promoting men who are here	2.5
You've got to have "pull" with certain people around here to get ahead	2.1
In the long run this company will "put it over" on you	1.5
The pay in this company is terrible	1.0
An honest man fails in this company	0.8

of the scale. The use of different statements in scales measuring the same attitude also facilitates checking of results by a repeat test in order to be sure of conclusions reached and to measure the effectiveness of systematic company efforts to improve employee morale.

Attitude scales are not perfect instruments for the registering of employee feelings, but they are considerably better than

guessing or relying on chance (and often biased) individual reports.

General employee opinion surveys

Attitude scales of the type described in the preceding section show the attitude of the employees toward the company as a whole. If it is possible to ask the employees to indicate the department in which they work without revealing their identity, such scales also enable a management to compare employee attitude from one department to another. But such scales do not enable a management to identify specific factors that may be sources of employee unrest or dissatisfaction. This information can be obtained by means of another type of questionnaire, which asks a number of specific questions concerning plant practices. This type of questionnaire should also be used without reference to any employee's name, and great care should be taken to insure the anonymous feature of the survey. An example of a questionnaire used in an employee survey of this type conducted by the Victor Adding Machine Company of Chicago is shown in Figure 133. The following steps were taken in conducting the survey employing this questionnaire:

1. Questions to be included were chosen by management and phrased so as to permit convenient and rapid recording of answers on electric punched cards for sorting and tabulating.
2. Supervision was informed of the project at special supervisory meetings.
3. Employees were informed of the project by supervisors. The employees were told that they would receive the questionnaires through the mail at their home addresses and that the company would appreciate their cooperation in checking the questionnaires and returning them to the tabulating agency (not the plant) in the enclosed self-addressed stamped envelope.
4. Questionnaires were received, results were tabulated, and a report of findings was prepared for management.

In the case of the survey conducted for this company, the procedure described above resulted in a 70 per cent return of

QUESTIONNAIRE

1. This questionnaire is being answered by:	Man ☐	Woman ☐
2. Are you—	Married ☐	Single ☐
3. Are you paid by the week or by the hour?	Paid by the week ☐	Paid by the hour ☐
4. Do you feel you would rather be doing some other type of work?	Yes ☐	No ☐
If yes, have you discussed it with the Personnel Office?	Yes ☐	No ☐
5. How do you regard safety conditions within the plant generally?	Good ☐	Not so good ☐
What is your opinion of your foreman or department head?		
6. Does he "know his stuff"?	Yes ☐	No ☐
7. Does he play favorites?	Yes ☐	No ☐
8. Does he keep you busy?	Yes ☐	No ☐
9. Does he keep his promises?	Yes ☐	No ☐
10. Does he pass the buck?	Yes ☐	No ☐
11. Does he welcome suggestions?	Yes ☐	No ☐
12. Is he a good teacher?	Yes ☐	No ☐
13. Do the workers know more than he does?	Yes ☐	No ☐
14. Does he set a good example?	Yes ☐	No ☐
15. Do you think you are in need of more training?	Yes ☐	No ☐
16. Which of the following ways of pay would you prefer?	Present way ☐	Incentive or piece work ☐
Do you feel you understand the following provisions of the Company Security Fund?		
17. The insurance provisions	Yes ☐	No ☐
18. How the Security Fund shares are figured	Yes ☐	No ☐
19. How VAMCO determines its contribution to the Fund each year	Yes ☐	No ☐
20. How the Security Fund money is invested	Yes ☐	No ☐
21. How your Fund account is closed in event of termination, death or retirement	Yes ☐	No ☐
22. Do you feel that you are receiving considerate treatment here?	Yes ☐	No ☐
If not, why____________________		
23. Do you feel top management is interested in the employes?	Yes ☐	No ☐
24. Have you ever recommended Victor as a place to work to a friend?	Yes ☐	No ☐
25. Would you like more news broadcasts given over the P. A.?	Yes ☐	No ☐
26. Are you interested in Company athletic activities?	Yes ☐	No ☐
27. Are you making suggestions to the Suggestion System?	Yes ☐	No ☐
28. Do you feel you have a good future with this Company?	Yes ☐	No ☐
29. Would you be interested in selling with Victor's post war Sales Division?	Yes ☐	No ☐
30. What is your reaction to being asked down to Personnel for occasional interviews?	Like it ☐	Don't like it ☐
31. Do you think Victor has done a good job for the country during the war?	Yes ☐	No ☐
32. Are you getting the kind of information about the Company that you want?	Yes ☐	No ☐

33. What do you think of working conditions here as compared with other plants?
Above average ☐ Average ☐ Below average ☐

34. How do you think your average weekly earnings (gross earnings before deductions) compare with that paid in other companys for the same type of work?
Better here ☐ About the same ☐ Lower here ☐

FIG. 133—Part of a questionnaire used in an employee survey conducted by the Victor Adding Machine Co. (*Courtesy of A. V. Larson, Asst. Personnel Director, Victor Adding Machine Co., Chicago, Illinois.*)

the questionnaires, a proportion far exceeding that obtained with many questionnaires sent out without previously notifying the employees that a questionnaire would be received.

A summary of the findings of the survey is not of particular importance to anyone not associated with the particular company and, therefore, is not given here. It is sufficient to say that the company made significant use of the results in planning the future activities of the industrial relations department.

The exit interview

Another method of determining both general employee attitude toward the company and also feeling or opinion with regard to specific practices, departments, or supervisors, is provided by the exit interview.

Onarheim,[14] summarizing the values of the exit interview, points out that it helps to retain desirable employees and provides a check on policies covering employment, placement, training, wages, reasons for terminations in different departments, and effectiveness of grievance procedure. Drake[15] also stresses the value of the exit interview in identifying sources of employee dissatisfaction.

An employee who is working for the company often is reluctant in expressing his true feelings. Also, many managements feel that it would not be good industrial relations practice to encourage present employees to give verbal expression to their feelings. But the employee who has terminated his employment usually is quite willing to say what he thinks about the company, and management has little to lose (and often much to gain) by listening attentively to what he has to say. This fact has resulted in an increasing use of the exit interview—an interview conducted with the employee leaving the company. While any single employee may have grievances for which there is no real basis in plant practices, yet if an appreciable number of employees terminat-

[14] J. I. Onarheim, "Exit Interviews Help Us Check Personnel Policies," *Factory Management and Maintenance*, CII (1944), pp. 121–122.

[15] C. A. Drake, "The Exit Interview as a Tool of Management," *Personnel*, XVIII (1942), pp. 346–350.

ing their employment mention the same situations or practices as unsatisfactory, it is usually safe to conclude that there is a real reason for their discontent. Under such circumstances, it is also a reasonable assumption that employees still on the job are not entirely satisfied with these practices or policies. Information obtained in a series of exit interviews should not, of course, be automatically followed by changes in managerial policy whenever a source of grievance is revealed, but most managements now believe that such information should at least be considered when making future managerial decisions.

Factors Affecting the Morale of Employees

A morale survey, the exit interview, or any other barometer of employee feelings, is of little practical value unless it succeeds in identifying factors that affect employee morale and points the way toward changes that may be instituted by management to improve morale. A number of investigations of this type have been conducted and several determining factors have been isolated. Some of these can be easily changed by management; others are more difficult to modify. In any event, a statement of certain factors that have been found related to employee morale will indicate problems faced by modern industrial relations departments.

Job or rank of employees

In Uhrbrock's study[16] it was found that foremen are more favorably disposed toward the company than are clerks, and clerks more favorably disposed than are factory workers. The distributions of attitude scores for the three groups are shown in Figure 134. This finding is in accord with what one might expect, and it may be considered as an added indication of the validity of the attitude-scale method of measurement. Many companies necessarily have some employees in jobs

[16] Uhrbrock, *op. cit.*

that lack status. The identification of such jobs and the means of improving morale of the employees assigned to them is one practical value of an attitude survey. Often a marked improvement of morale can be effected by a slight change in job status. Super[17] has found that the *amount* of change in status is of little importance in effecting an increase in job satisfaction, but the *direction* of change is of vital importance. Those departments having employees whose morale would be helped by such measures can be readily

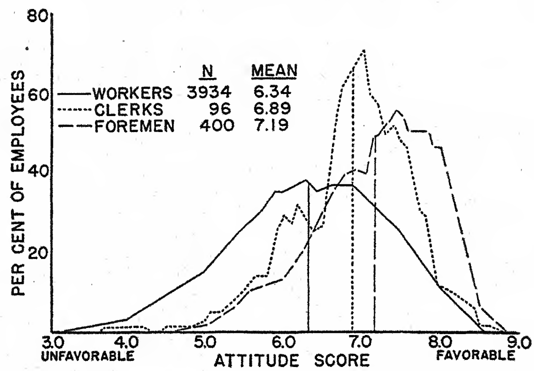


FIG. 134—Distribution of attitude toward the company of three groups of employees.

identified by means of an anonymous attitude survey, provided only that the employees identify their department and job.

Success of the employee

The old maxim that "nothing succeeds like success" applies to employee morale. Those employees who are doing a job well and are made to feel that they are doing a job well have good morale. If they are failing, or made to feel that they are failing, morale suffers. For high morale to exist, the job does not have to be one of vital importance to

[17] D. E. Super, "Occupational Level and Job Satisfaction," *Journal of Applied Psychology*, XXIII (1939), pp. 547–564.

the plant. It can be quite an ordinary job and yet accomplish the desired result. In other words, it is not so much the job as how the employee feels about it and how the boss regards the employee that determines the latter's morale. Glen Gardiner[18] has given an excellent illustration of this principle:

> Riding on a freight elevator with the superintendent of the U. S. Gypsum Company on Staten Island, I was impressed by this incident. As we stepped on the elevator the superintendent said, "I want to introduce you to Tony who has been running this elevator for more than eight years and has never had a single accident of any kind, which is a record we're proud of."
>
> The way Tony grinned and stuck out his hand made one realize that he appreciated the credit being given him and that he was 100 per cent sold on the idea of running that elevator for many years to come without any accidents.

Comprehensive statistical studies have been made which show that the morale of the successful employee is better than that of the employee who is struggling with his job. For example, Kolstad[19] found that the morale scores of low-selling employees in a department store were significantly lower than the corresponding scores of the high-selling employees. It seems clear that employees should not be placed nor kept on jobs in which they are unable to achieve a reasonable degree of success.

Supervision

Everyone recognizes the fact that a man may work contentedly for one boss and complain bitterly when asked to work for another. Probably few things exert so unfavorable an effect on employee morale as the blustering, bully type of "straw boss." The importance of the supervisor has been emphasized by the work of McMurry[20] and also by that of

---

[18] Glen Gardiner, "Reaching the Individual Worker," Address before the Second Annual Greater Philadelphia Safety Congress, May 27, 1935.

[19] A. Kolstad, "Employee Attitudes in a Department Store," *Journal of Applied Psychology*, XXII (1938), pp. 470–479.

[20] R. N. McMurry, "So You Handle People," *Society for the Advancement of Management Journal*, II (1937), pp. 168–172.

Hull and Kolstad.[21] In the past, management has too often selected the supervisor from the working group primarily because the man chosen was a good worker or a high producer. We now know that ability to produce well on that job is no guarantee whatever of ability to supervise other men on the job. The supervisor chosen because of his production record is as likely as not to be a failure in handling men; and we might add that he will almost certainly fail unless he is given specific training in how to handle men. Industry today recognizes this situation and is not only considering many factors besides production ability in promoting a man to a supervisory position, but also is training supervisors in the solution of problems unique to the supervisory job. Evidence shows that when supervisors and others in positions of authority are correctly trained, the guidance which they may give has a favorable effect on employee morale.[22]

Social factors in the working group

The importance of group solidarity in the small groups of workers that almost automatically are formed in any plant has not been very well recognized by many managements. Yet the work of Mayo[23] and his associates clearly demonstrates the marked relation between feelings of "belonging" to the group and the general level of employee morale. In one study reported, of "mule spinners" in a textile mill, not only did increased production and reduced labor turnover result after steps were taken to transform a horde of "solitaries" to a group with a group-consciousness, but also the general social changes—even outside the plant—were in a direction indicating a higher level of general job and social satisfaction. Similar results were found in the Hawthorne study. In commenting upon the results of these studies Mayo writes:

---

[21] R. L. Hull and A. Kolstad, "Morale on the Job," in G. Watson, *Civilian Morale* (Houghton Mifflin, 1942), pp. 349–364.

[22] J. F. Murphy, O. M. Hall, and G. L. Bergen, "Does Guidance Change Attitudes?" *Occupations*, XIV (1936), pp. 949–952.

[23] Mayo, *op. cit.*

"For all of us the feeling of security and certainty derives always from assured membership of a group. If this is lost, no monetary gain, no job guarantee, can be sufficient compensation. Where groups change ceaselessly as jobs and mechanical processes change, the individual inevitably experiences a sense of void, of emptiness, where his fathers knew the joy and comradeship and security."[24] Perhaps one of the reasons that rest periods during the working shifts are usually accompanied by improved job performance is that during rest periods the workers have an opportunity to form and solidify a group consciousness that cannot take shape while every operator is occupied at his machine. In the illustration of the "mule spinners" referred to, the introduction of four ten-minute rest periods during the day was the major change made by management. There was a general feeling by supervision that there would be no way to "make up" the lost forty minutes. Yet when the group consciousness had been formed, job performance improved and there was general agreement that the time was indeed not "lost" at all.

Working conditions

Employee morale is lowered by unfavorable working conditions. This situation is the more serious because often the employees who are affected are unaware of the true cause of their grouchiness and dissatisfaction. Collier[25] has pointed out a number of specific relationships found between reactions of employees and the physical working conditions of several jobs. He states that the feelings of uneasiness and unrest found among spray painters were removed by the installation of a new air exhaust. He also found that the presence of methylene chloride produces bad temper, irritability, and sleeplessness. Needless to say, the employees so affected were unaware of the real cause of their mental discomfort

[24] Mayo, *op. cit.*, p. 76.

[25] H. E. Collier, "The Mental Manifestations of Some Industrial Illnesses," *Occupational Psychology*, XIII (1939), pp. 89–97.

and therefore probably tended to attribute their feelings to imaginary causes. Other effects reported by Collier were that mercury produces irritability, anxiety, depression, and sleeplessness, while manganese produces languor, lethargy, muscular cramps, and abnormalities of gait.

These findings indicate that the physical surroundings of a job may result in a general lowering of employee morale in a way which is not revealed by ordinary questioning of the men. The conclusion we may draw is that when an unfavorable morale condition is found to exist, it is wise to examine carefully the physical surroundings to determine whether some unnatural condition may be at fault.

Some evidence indicates that a deliberate effort to create favorable working conditions (as opposed to the elimination of definitely unfavorable surroundings) will be followed by an improvement of morale. For example, Tindall[26] reports that the presence of music on the job speeds up production, improves morale, pacifies labor unrest, creates good will, lessens labor turnover, and reduces errors. Kerr[27] has reported a series of studies showing that under certain conditions music has a favorable effect upon the attitude of employees toward their job. He also reports that ". . . when other factors are equal, workers will go to the job locations where they can hear music while they work."[28] While the several favorable effects found in these studies suggest the desirability of keeping the working conditions as pleasant as the nature of the work will permit, we should hardly conclude that the use of music is either possible or desirable in all instances. Nor should we assume that the use of music will automatically improve the production or

---

[26] G. M. Tindall, "Rhythm for the Restless," *Personnel Journal*, XVI (1937), pp. 120–124.

[27] W. A. Kerr, "Psychological Effects of Music as Reported by 162 Defense Trainees," *Psychological Record*, V (1942), pp. 205–212.

[28] W. A. Kerr, "Where They Like to Work: Work Place Preference of 228 Electrical Workers in Terms of Music," *Journal of Applied Psychology*, XXVII (1943), pp. 438–442.

quality of work. As Kirkpatrick[29] points out "no highly significant or conclusive research has been published concerning the effect of music on output or health of workers in industry." And, in many cases (for example in a steel mill) music would be out of the question, because the general noise level is so great that no one could hear it.

Salary reviews and praise

A fact not realized by many hopeful college graduates is that it is entirely possible for an employee to become lost and buried in a large industrial plant, *even while he is doing his job well.* Indeed, the very fact that he is doing a routine job well may prevent him from being brought to the attention of anyone who has the authority to promote or demote him. Obviously, the permanent or semi-permanent location of an employee on a job from which he hoped to advance is not conducive to high morale. Many industries are solving this problem by a scheme for systematically reviewing, at stated intervals, the job and earnings of every employee. Such a scheme does a great deal to prevent a capable employee from being "lost in the mill." The advisability of such a scheme has been stressed by Shepard[30] in his emphasis of the fact that workers are happier as well as more valuable if they are praised and given regular salary reviews.

The matter of praise, also, is sometimes given too little consideration by management and supervision. If we are willing, temporarily, to oversimplify a very complex problem, we might say that there are in general two ways of motivating people: to praise (or reward) for doing the desired thing, and to reprove (or punish) for doing the wrong thing. Much thought and considerable experimentation have been concentrated on this problem in an effort to determine which of these methods of motivation achieves the best results.

---

[29] F. H. Kirkpatrick, "Music in Industry," *Journal of Applied Psychology,* XXVII (1943), pp. 268–274.

[30] J. L. Shepard, "Recognition on the Job," *Personnel Journal,* XVI (1937), pp. 111–119.

Unfortunately, a general conclusion that is universally applicable in all situations has not been reached; but the preponderance of evidence clearly favors the praise over the reproof method. This was the conclusion reached by Hurlock[31] in her studies of learning in school children, and there seems to be every reason for expecting the results to apply in a general way to the behavior of industrial employees. One should not infer from Hurlock's study that praise is always superior to reproof or that reproof should never be used. Her investigations, indeed, show that reproof for failure accomplishes better results than no comment at all; and in practical situations, it seems obvious that situations arise which clearly call for some kind of reproof. The general principle that we emphasize here is that praise and rewards are in many (if not most) cases more potent than reproof as a motivating factor and will *almost always have a better effect on employee morale.* Supervisors will do well not only to *know* this, but to *use* it in their day-by-day operations with their men.

Wage payment methods

Entirely aside from the *amount* of money earned by an employee, the *method* by which his wage is computed may be a source of employee complaint. Wage incentive plans are ordinarily installed for the purpose of motivating the worker to reach a reasonable output and at the same time of rewarding the employees who reach such an output in proportion to the actual amount of work done. A discussion of the various wage incentive plans in use by modern industry does not fall within the scope of this volume; but it should be said here that such plans are often somewhat complicated in operation and are therefore not always understood completely by the employees who are affected by them. An employee wants to know how to figure his own check. If no one has explained the wage payment plan to him, or if it has been explained but

[31] E. B. Hurlock, "An Evaluation of Certain Incentives Used in School Work," *Journal of Educational Psychology*, XVI (1925), pp. 145–159.

he has not understood the explanation, he is very likely to be suspicious of the whole system. An employee who figures that the company owes him $48.60 but who receives a check for only $48.50 is disgruntled out of all proportion to his imagined loss of ten cents. When management installs a wage incentive plan, the job is only half done; there remains the job of *explaining* the plan to the employees who are to work under it, of being sure that they *understand* it, and of *selling* it to them. Modern management no longer tells its employees, "This is the way we are going to do it, and you can like it or quit." Management today realizes that it is good business to help every employee to feel that he is part of the business. Such a feeling cannot be developed if an employee cannot even compute his own pay check.

The same precautions should be observed when hourly rates are paid under a job evaluation system and wage structure. Job evaluations (see page 374) are based upon a job analysis of the various jobs so that equal skill, training, and effort on the different jobs will be paid equally well. Such job analyses always involve some subjective element in rating the different jobs; and it is a rare case where the first installation of a job evaluation system is not found to contain a few jobs that are "out of line," that is, are set too high or too low in comparison with other jobs. It is very important to keep a constant lookout for such cases and to correct them at the first opportunity. Both in factory work and office work, jobs that are "out of line" in rate or salary are a significant source of employee unrest.[32]

Other factors affecting morale

Several other studies have identified still additional factors which, at least in certain cases, have an effect on employee

[32] H. B. Bergen, C. E. Haines, L. G. Giberson, F. L. Hallock, and C. S. Coler, "Attitudes and Emotional Problems of Office Employees," *Office Management Series*, No. 87, p. 34.

P. Hall and H. W. Locke, *Incentives and Contentment: A Study Made in a British Factory* (Pitman, London, 1938).

morale. In Uhrbrock's investigation it was found that the attitude toward the company of the male employees was significantly less favorable than that of the female employees. Apparently some factor, or combination of factors, which had been unnoticed by management was operating in this particular company to give the women a better feeling toward the company. The forces at work may have been seemingly trivial to *management*, but factors that seem trivial to management often seem quite important to employees. In fact, as Schultz[33] has pointed out, employee attitudes are often affected more by little things than by broad management policies.

It should be mentioned also that employee attitudes are often made up of a combination of attitudes toward different aspects of the total situation. Employees may be favorable toward some, and unfavorable toward other, aspects of their jobs. Geiger, Remmers, and Greenly[34] found that little relationship exists between the attitude of apprentices toward their job proper, fellow employees, foremen, opportunity for promotion, related instruction, and management. This study includes the development of a diagnostic scale which indicates how differential attitude toward different aspects of the job may be measured.

It is not intended to imply that the factors listed above are *all* of the factors that may affect employee morale. The list is only illustrative of factors that have been found operating in certain plants and that may give cues to the alert plant superintendent in analyzing his own particular situation. Although we may confidently expect that future developments in applied psychology will make this problem easier to cope with than it is today, sufficient evidence already exists to show that the judicious use of attitude scales and of techniques

[33] R. S. Schultz, "Psychology in Industry," *Personnel Journal*, XVI (1937), pp. 221–223.

[34] H. E. Geiger, H. H. Remmers, and R. J. Greenly, "Apprentices' Attitudes Toward Their Training and the Construction of a Diagnostic Scale," *Journal of Applied Psychology*, XXII (1938), pp. 32–41.

now available will aid in the solution of many industrial relations problems. Departments in which unsatisfactory relations exist between employees and their supervisor can be located. Company policies that are not satisfactory to employees can be identified. Employee reaction to such topics as method of wage payment, insurance programs, and plans for promotion and transfer can be determined. The employee attitude survey and the exit interview are of real value in the solution of such problems and have become a powerful tool of management in the prevention of industrial relations disputes.

Appendices

Appendix A

Elementary Statistical Procedures

WHEN many measurements, such as scores or other data, are to be summarized or interpreted, the use of some form of statistical procedure is usually desirable. If a considerable amount of raw data is involved, a simple listing of the data is of little value. Such a listing will not tell us, for example, how the data are distributed, how much they vary, or where in the total distribution they tend to cluster. Further, such a listing is of little value in indicating how the data compare with, or are related to, other sets of data collected under other circumstances. Before a meaningful interpretation of the data can be made it is necessary to reduce them to a *chart or to one or two single numbers which may represent the data as a whole.*

Graphic Representation of Data

The frequency distribution and polygon

A frequency polygon, constructed from a frequency distribution, is a graphic representation of a set of data. The construction and interpretation of a frequency polygon may best be explained by an example. Suppose 60 employees on an inspection job have detected the following number of flaws of a certain type during one week of work:

TABLE 55

NUMBER OF FLAWS DETECTED BY EACH OF 60 INSPECTORS DURING ONE WEEK OF WORK

15	36	40	37	32	13	35	20	33	36	33	16	38	19	33	34	24
36	25	29	27	39	42	31	21	26	28	53	23	51	21	26	39	28
30	31	32	30	29	49	39	30	44	34	37	35	38	35	41	37	43
42	38	45	22	46	41	47	48	34								

From a gross, or even a detailed, inspection of these 60 values one cannot answer such questions as: What is the typical number of defects spotted by an average inspector in a week? How much difference is there between the best and poorest inspectors in spotting defective material? Is there any preponderance of good,

poor, or average ability represented in the performance of these inspectors?

Such questions as these may be answered at a glance if the data are grouped and presented in a chart. One variety of such a chart is a frequency polygon. The steps involved in constructing a frequency polygon are as follows:

1. Determine the range of the values in the raw data. Quickly glance through the data to determine the *highest* and the *lowest* values. The range is the difference between these values. In the case of the 60 inspector records, the highest figure is 53 and the lowest is 13. The range is therefore 53 − 13 = 40.

2. If we find that the range of the data is large (that they are widely spread), it will be more convenient to group them by intervals (class intervals, abbreviated c.i.) with a range in each c.i. of more than one unit. The c.i. is a group of adjacent scores of such a size that from 12 to 18 c.i.'s cover the range of the whole distribution or all of the data. With a range of 40, a c.i. of 2 would require 20 groups; a c.i. of 3, 14 groups; and a c.i. of 4, 10 groups. In our illustrative problem, a c.i. of 3 is therefore the proper size to use.

A simple rule of thumb which is helpful in deciding upon the correct size of the c.i. is to divide the range by 15 (15 because, on the average, this is the most desirable number of c.i.'s) and take as the c.i. the whole number nearest to the quotient. In our problem, the range divided by 15 would be 40 ÷ 15 = 2.66. As 3 is the whole number nearest to 2.66, 3 would be the size of the c.i. to be used.

3. Arrange the adjacent c.i.'s in a column, leaving a blank space immediately to the right of this column. The arrangements of the c.i.'s preparatory to the construction of a frequency distribution appears as follows.

TABLE 56

CLASS INTERVALS TO BE USED FOR ILLUSTRATIVE DATA IN TABLE 55

51–53
48–50
45–47
42–44
39–41
36–38
33–35
30–32
27–29
24–26
21–23
18–20
15–17
12–14

4. Place a tally mark for each value in the original list of raw data opposite the appropriate class interval. As the first value among the 60 listed in Table 55 is 15, the first tally mark should be in the 15–17 c.i. The second value, 36, is represented by a tally mark in the 36–38 c.i. Usually it is advisable to tally the fifth entry in each c.i. with a line across the preceding four tally marks. This simplifies the counting of tally marks at a later time. When all entries have been made, that is, all data tabulated, the frequency distribution appears as in Table 57.

5. Lay off appropriate units on squared (cross section or graph) paper so that a graph may be constructed on which the midpoints of

TABLE 57
Class Intervals, Tally Marks, and Frequencies (f.) for Illustrative Data in Table 55

Class Intervals (c. i.)	*Tally Marks*	*Frequency (f.)*
51–53	//	2
48–50	//	2
45–47	///	3
42–44	////	4
39–41	~~////~~ /	6
36–38	~~////~~ ////	9
33–35	~~////~~ ////	9
30–32	~~////~~ //	7
27–29	~~////~~	5
24–26	////	4
21–23	////	4
18–20	//	2
15–17	//	2
12–14	/	1
	Total =	60

the c.i.'s are plotted on the base line and the frequencies or number of cases in each c.i. on the vertical axis. When this is done, the frequency polygon shown in Figure 135 is obtained.

To one familiar with the concept of a frequency polygon, the graphic illustration in Figure 135 is a much more meaningful representation of the data than the list of values shown in Table 55, or the frequency distribution shown in Table 57. The frequency polygon makes apparent at a glance that the typical or average inspector detected around 35 defects during a week of work, that the operators range or vary from some who detected only 13 defects to others who detected 52, and that a majority of the operators are fairly near the average in ability (that is, that not so many are very high or very low as are near the average). In summarizing psy-

chological data it is a definite advantage to be able to present all of the major facts in a single graphic presentation of this type.

In a frequency polygon such as the one shown in Figure 135, the area between the curve and the base line is determined by the number

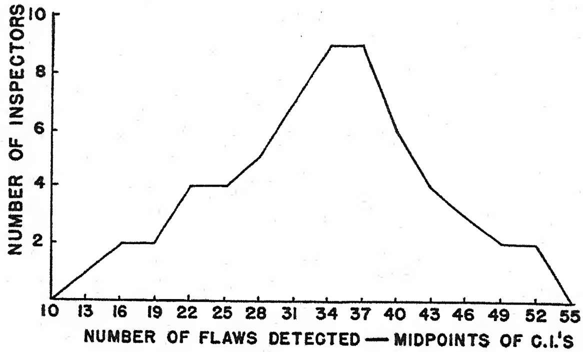


FIG. 135—Frequency polygon of illustrative data in Table 57.

of cases (called N) which the graph represents. Thus, a curve portraying 120 cases would cover twice the area of the curve shown (if the c.i.'s are the same for both distributions), and a curve

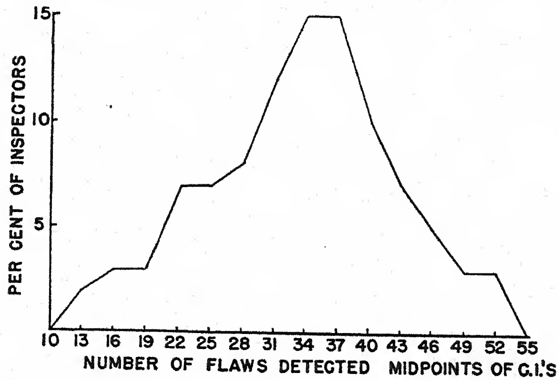


FIG. 136—Frequency polygon with ordinates as percentages for illustrative data in Table 58.

portraying 600 cases would cover ten times this area. This is no disadvantage in many cases, but situations sometimes arise in which it is desirable to keep the total area under the curve the same, regardless of N or the number of cases. To meet this situation we

may plot the *percentage* of cases rather than the *number* of cases falling in each c.i. This may be accomplished by dividing each f value in Table 57 by the total number of cases in the distribution. Each quotient thus obtained indicates the percentage of cases from the total falling in the respective c.i. These computations are indicated in Table 58.

If the percentages shown in the last column of Table 58 are now plotted as the ordinates (vertical axis points) of a frequency polygon, the chart shown in Figure 136 is obtained.

When a frequency polygon is to be compared with a number of other polygons, and when the important facts to be compared deal with the central tendencies and general form of the distributions rather than with the different number of cases plotted in each, the

TABLE 58
FREQUENCY AND PERCENTAGE OF CASES IN EACH C.I. FOR ILLUSTRATIVE DATA IN TABLE 55

Class Intervals	f	*Calculation of Percentage*	*Percentage*
51–53	2	$\frac{2}{60} = .033$	3
48–50	2	$\frac{2}{60} = .033$	3
45–47	3	$\frac{3}{60} = .050$	5
42–44	4	$\frac{4}{60} = .066$	7
39–41	6	$\frac{6}{60} = .100$	10
36–38	9	$\frac{9}{60} = .150$	15
33–35	9	$\frac{9}{60} = .150$	15
30–32	7	$\frac{7}{60} = .117$	12
27–29	5	$\frac{5}{60} = .083$	8
24–26	4	$\frac{4}{60} = .066$	7
21–23	4	$\frac{4}{60} = .066$	7
18–20	2	$\frac{2}{60} = .033$	3
15–17	2	$\frac{2}{60} = .033$	3
12–14	1	$\frac{1}{60} = .017$	2
Total =	60		100

"percentage method" of plotting a frequency distribution is preferred to the "total-number-of-cases method."

We may define a frequency polygon, in the light of the above description, as *a curve which portrays data graphically and which is so drawn that the base line represents the varying values of the original data and the ordinates represent the number of cases (or percentage of cases) at each of the raw data values.*

The normal distribution

The shape of the frequency polygons shown in Figures 135 and 136 is typical of the kind of distribution usually found when data obtained from a group of people are plotted. It will be noted that the curves are approximately "bell-shaped," that is, they are high

in the center and taper off toward the base line at both ends. If we were to divide the area under such a curve by drawing a perpendicular line from the central high point to the base line, the two parts would be approximately equal in area and would be bilaterally symmetrical in shape. It is well recognized that all, or nearly all, measurements of human traits and abilities result in distributions of approximately this form. Such distributions are called *normal distributions*. A strictly normal distribution conforms to a symmetrical bell-shaped curve that is defined by a mathematical equation, the derivation of which is beyond the scope of the present discussion.[1] It will suffice for the beginning student to know that:

1. A normal distribution is bell-shaped, that is, it is high in the center and low at both ends. Its two halves are symmetrical.
2. Measurements obtained from a group of persons usually approximate this type of distribution.

Measures of Central Tendency

While the frequency polygon is helpful in giving an immediate graphic description of a set of data, that is, indicating general trends, it is often desirable to present certain quantitative figures that supplement the graphic picture. One of the most useful of such numerical values is a measure of the central tendency of the data. A measure of central tendency may be defined as a single figure or value which is representative of the entire set of data. Three such measures that are in common use are the arithmetic mean, the median, and the mode.

The arithmetic mean

The arithmetic mean, sometimes simply called the mean, may be defined as the sum of the measures divided by the number of measures.[2] Or it may be thought of as a point of balance which could be found if all values in the distribution were assigned the same weight and then arranged along a horizontal beam. The physicist might define it as that point in the distribution around which the moments are equal.

In the case of the 60 values previously discussed from which a

---

[1] The interested student is referred to C. C. Peters and W. R. Van Voorhis, *Statistical Procedures and Their Mathematical Bases* (McGraw-Hill, New York, 1941), pp. 270–286.

[2] H. E. Garrett, *Statistics in Psychology and Education*, 3rd ed., (Longmans, Green and Company, 1947).

frequency polygon was constructed, the mean is obtained by finding the total of the 60 measures and dividing this total by 60, thus:

$$\text{Arithmetic Mean } (A.M.) = \frac{\text{Sum of measures}}{N} = \frac{2016}{60} = 33.6$$

This is the procedure followed in computing the exact value of the arithmetic mean of any set of values. In practice a shorter method of computation utilizing the data as tabulated in a frequency distribution and yielding an approximation (rather than the exact value) of the mean is often used. This shorter method assumes that each score as tabulated in a frequency distribution has the same value as the midpoint of the c.i. in which it falls. For further convenience in calculation, the mean is first computed in c.i. units from an arbitrary base selected near the center of the distribution at the midpoint of one of the c.i.'s. The base selected is entirely arbitrary—it may be taken as any point in the distribution. We have chosen one near the center of the distribution to simplify computation.

If this method is applied to the frequency distribution in Table 57 the arrangement shown in Table 59 is obtained.

In the above tabulation the *d* column represents the number of c.i. units each c.i. is located above or below the c.i. arbitrarily chosen as the base for calculations. In the c.i. 51–53 there are 2 scores. This c.i. is 6 c.i. units above the arbitrary base. Thus,

TABLE 59

Computation of the Arithmetic Mean from a Frequency Distribution

c.i.	*f.*	*d.*	*fd.*
51–53	2	6	12
48–50	2	5	10
45–47	3	4	12
42–44	4	3	12
39–41	6	2	12
36–38	9	1	9
33–35	9	0	0
30–32	7	−1	−7
27–29	5	−2	−10
24–26	4	−3	−12
21–23	4	−4	−16
18–20	2	−5	−10
15–17	2	−6	−12
12–14	1	−7	−7
	60	$\Sigma fd =$	−7

Formula for Computing *A.M.*

$$A.M. = M^\circ + c.i. \frac{\Sigma fd}{N}$$

M° = assumed mean

$c.i.$ = size of *c.i.*

$c = \frac{\Sigma fd}{N}$ = summation of deviations from assumed mean divided by N

$$A.M. = 34 + 3\left(\frac{-7}{60}\right)$$
$$= 34 - .35 = 33.65$$

in computing the A.M. in c.i. units from the arbitrary base, these two scores would each have a value of 6, resulting in the number 12

which appears in the fourth, or *fd*, column. In like manner, there are 2 scores in the c.i. 48–50, and these two scores are each 5 c.i. units above the arbitrary base, resulting in the number 10 which appears in the *fd* column. All scores tabulated in c.i.'s below the arbitrary base are represented by negative values in the *fd* column. The algebraic sum of this column (Σfd) divided by the number of cases indicates how far the computed mean will deviate from the assumed mean (base) in terms of c.i. units. From the tabulation, this deviation in c.i. units from the arbitrary base (assumed mean) is defined as:

$$\text{Deviation in c.i. units from base} = \frac{\Sigma fd}{N}$$

Carrying through this computation for the data under consideration shows that:

$$\text{Deviation in c.i. units from base} = \frac{\Sigma fd}{N} = \frac{-7}{60} = -.117$$

This is interpreted to mean that the *A.M.* is .117 of a class interval below the midpoint of the arbitrary base (see formula in illustrative problem). In order to transmute this deviation (−.117) into raw score units we would multiply it by 3 (the size of the class interval). Thus, in terms of raw score units, the deviation is −.35. The mean, as computed by this method, is therefore .35 raw score units below the midpoint of the 33–35 c.i. As the midpoint of this is 34, the mean is 34 − .35 = 33.65. This approximation does not agree exactly with the exact method in which all raw data are added and the sum is divided by the number of cases; but the approximation is sufficiently close to justify its use in many cases. The student may note, however, that essentially the same procedures are used in both solutions. The procedure in using the "exact method" may be thought of as involving the computation of a mean by assuming the mean to be zero, computing the deviations from zero in raw score units, and dividing their sum by N as in the short method.

The median

The median is a measure of central tendency defined as that score (or value) which exceeds, and is exceeded by, half the measures, that is, it is that point in the distribution above and below which 50 per cent of the values lie. A logical (though laborious) method to determine the median consists in arranging all the raw data in rank order from lowest to highest and counting off the bottom half of the measures. The value at this point is the median.

If this method is followed for the data in Table 55, the following arrangement of the scores is obtained.

TABLE 60

Arrangement of Data for the Computation of the Median Directly from Raw Data

53	45	41	38	36	34	32	29	26	21
51	44	40	38	36	34	31	29	25	20
49	43	39	37	35	33	31	28	24	19
48	42	39	37	35	33	30	28	23	16
47	42	39	37	35	33	30	27	22	15
46	41	38	36	34	32	30	26	21	13

Counting from the lowest score up, we find that the 30th from the low end is 34, and the 31st from the low end is also 34. The median score would therefore be 34. If there had been a difference between the 30th and 31st scores, the median would be the value halfway between these scores. If an odd number of cases were included in the original set of scores (as 61 instead of 60) the median would be the value of the middle score.

In practice, the median as well as the mean may be conveniently approximated from a tabulated frequency distribution. To illustrate this process we may use the same frequency distribution previously discussed.

TABLE 61

Computation of the Median from a Frequency Distribution

Class Intervals	*f*	*Cumulative f*
51–53	2	60
48–50	2	58
45–47	3	56
42–44	4	53
39–41	6	49
36–38	9	43
33–35	9	34
30–32	7	25
27–29	5	18
24–26	4	13
21–23	4	9
18–20	2	5
15–17	2	3
12–14	1	1
Total	60	

Formula for Median

$$Md. = l + c.i. \frac{\left(\frac{N}{2} - F\right)}{f_m}$$

l = lower limit of c.i. within which median lies.

$c.i.$ = class interval

$\frac{N}{2} = \frac{1}{2}$ of the scores

F = no. of scores in all $c.i.$'s below l

f_m = no. of scores in $c.i.$ in which median falls.

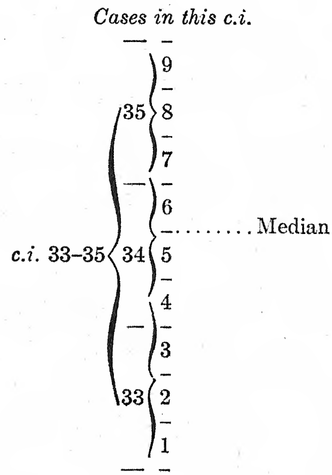
$$Md. = 32.5 + 3\,\frac{(\frac{60}{2} - 25)}{9}$$

$$= 32.5 + 1.7 = 34.2$$

Knowing that 60 cases are included in the distribution, it is necessary to find the score which separates the lower 30 from the upper 30. The value of this score is the median. Counting up from the lower part of the distribution, we first fill the column

cumulative f, which indicates for each c.i. the number of cases *in and below that c.i.*

It will be noted that 25 cases are included in or below the 30–32 c.i., and that 34 cases are included in or below the 33–35 c.i. The median, or point midway between the 30th and 31st case, must therefore be within the 33–35 c.i. Now if we assume that the 9 cases in the 33–35 c.i. are distributed evenly throughout this interval, we must go up into this c.i. far enough to cover the lowest 5 of these 9 cases in order to reach the median. This may be illustrated graphically as follows:



It will be noted that the "real limits" of the 33–35 c.i. are considered 32.5 and 35.5, that is that the c.i. extends one-half a score unit above and one-half a score unit below the tabulated values. This is necessary because otherwise there would be a whole unit between each pair of adjacent c.i.'s that does not logically fall in either c.i.

In this case, the median would be $\frac{5}{9}$ of the size of the c.i., added to the lower limit of the c.i., or Median $= \frac{5}{9}\,(3) + 32.5 = 34.2$. This value, 34.2, differs slightly from the value computed by arranging the scores in rank order, but the approximation is sufficiently close to justify its use in most instances. Exactly the same procedure can be followed in defining other points in the distribution, for example, the 25th percentile—the point in the distribution below which 25 per cent of the scores lie and above which 75 per cent lie.

The mode

A third measure of central tendency is the mode, which is defined as the measure appearing most frequently. This value, as well as the mean and the median, may be determined directly from the raw data (if one value appears more often than any other) or may be approximated from a frequency distribution of the data.

In computing the mode directly from the raw data, the values are inspected to determine which one appears most frequently. Sometimes, as in the case of the values shown in Table 55, several of the measures appear an equal number of times. In this case, the figures 30, 33, 34, 35, 36, 37, 38, 39 each appear three times. It is incorrect, therefore, to say that any one of these is the mode. Furthermore, there is reason to believe that if a larger sample than 60 inspectors had been included in the distribution, and if the c.i. used in forming the distribution were smaller than 3, the frequency polygon obtained would be more even in curvature and only one single high point would be found. Under such circumstances, this high point would be the mode. An approximation of this value may be obtained from a frequency distribution by means of the following empirical formula:

$$\text{Mode} = 3 \text{ Median} - 2 \text{ Mean}$$

In the case of the data we have been discussing, this formula gives the following value for the mode:

$$\text{Mode} = 3(34.2) - 2(33.65) = 35.30.$$

When to use the mean, median, and mode

Why is it necessary to have three different measures to indicate the central tendency of a set of data? The answer is that each is best adapted to certain uses, that is, in some cases one may be most representative of a set of data, while in other cases, another measure may be most suitable. The mean is ordinarily used if the distribution is approximately normal. (If the distribution is perfectly normal, the three measures of central tendency have the same value.) If, on the other hand, there is a preponderance of extreme cases at either end of the distribution, the mean may give an incorrect impression of the central tendency of the data. Under these circumstances, the median or mode is more suitable. Consider, for example, the following yearly incomes of five persons:

$800 $900 $850 $750 $5000

The mean for these five incomes is

$$\frac{\$800 + \$900 + \$850 + \$750 + \$5000}{5} = \$1660.$$

This figure, though an accurate statement of the mean, is not typical of the group as a whole because it is so markedly affected by the one income of \$5000 which is considerably larger than the other four. The median income is \$850, and this value is more typical for the group as a whole than is the mean income of \$1660. If a great deal of data were available for computation, it would also be enlightening to know the mode, or most common income. Certain other principles also help determine which measure of central tendency is most appropriate in any specific case. We may generalize the above illustration by saying that if a distribution is very much *skewed* (that is, contains more cases at one extreme than at the other), the median or mode is more likely to give a representative picture of the typical score than is the mean.

Measures of Variability

In addition to a measure or value to represent the central tendency of a set of data, there is also quite frequently a need for some measure of the spread, or variability, of the data. The need for a measurement of this type may be seen by comparing the data shown in Table 55 and tabulated in Table 59 (the mean of which, computed from the frequency distribution, was found to be 33.65) with another set of data which, for purposes of illustration, we might assume to consist of 21 scores of 33, and 39 scores of 34, making 60 scores in all. The mean of 60 such scores may readily be found to be 33.65.

$$\frac{(21)(33) + (39)(34)}{60} = 33.65.$$

While both distributions have the same mean, they differ markedly in variability or spread. The former distribution is made up of scores varying from 13 to 53, while the latter consists entirely of scores of 33 and 34. A quantitative measure of variability is therefore of considerable value. Statistical procedures have been designed which yield a single value descriptive of this variability; as in the cases of means and medians these measures tell us something about the group as a whole. Two measures of variability of a set of scores or other data are the Mean (or Average) Deviation and the Standard Deviation.

The mean or average deviation

This measure of variability is defined as the average deviation of the scores from the central tendency, usually the arithmetic mean, but sometimes the median or mode. If the arithmetic mean is used as the central point from which the mean deviation is computed, the computation involved for the 60 scores tabulated in Table 59 would be as follows:

In Table 62 columns 1–4 are identical with the corresponding columns in Table 59. Column 5 gives for each c.i. the deviation

TABLE 62

COMPUTATION OF THE AVERAGE DEVIATION (A.D.) FROM A FREQUENCY DISTRIBUTION

c.i.	*f*	*d*	*fd*	D^* *Raw Score Deviation from Mean*	*fD*
51–53	2	6	12	18.35	36.70
48–50	2	5	10	15.35	30.70
45–47	3	4	12	12.35	37.05
42–44	4	3	12	9.35	37.40
39–41	6	2	12	6.35	38.10
36–38	9	1	9	3.35	30.15
33–35	9	0	0	.35	3.15
30–32	7	−1	−7	2.65	18.55
27–29	5	−2	−10	5.65	28.25
24–26	4	−3	−12	8.65	34.60
21–23	4	−4	−16	11.65	46.60
18–20	3	−5	−10	14.65	43.95
15–17	2	−6	−12	17.65	35.30
12–14	1	−7	−7	20.65	20.65
			$\Sigma fd = -7$		441.15

$$\text{Mean} = 34 - (\tfrac{7}{60})3 = 33.65$$

$$\text{Mean dev.} = \frac{441.15}{60} = 7.35$$

* Note that all deviations, whether above or below the mean, are taken as positive.

between the midpoint of the interval and the arithmetic mean of the distribution. For the c.i. 51–53 this deviation is 52–33.65 = 18.35, which appears as the first value in the fifth column. Since there were two scores in this c.i., and since each deviates by 18.35 from the mean of the distribution, the *fD* column contains 2(18.35) = 36.70 as the first entry. In like manner, the amounts of deviation of all scores in the remaining c.i.'s have been computed and entered in the *fD* column. The sum of this column is the total of the deviations of all 60 scores, and this total divided by 60 gives the Mean Deviation of the distribution.

Had the 60 scores been bunched around the mean to a greater extent, the Mean Deviation would necessarily have been smaller. If, for example, all 60 scores had been exactly *at* the mean, the deviation between the mean and each of the scores would have been zero, the sum of the 60 deviations would be zero, and the Mean Deviation would be zero. In the illustrative case previously referred to consisting of 60 scores made up of 21 scores at 33 and 39 scores at 34, the mean deviation, though not zero, would be very small. In this case, the Mean Deviation may be computed in the following fashion: the 21 scores at 33 contribute 21(.65) = 13.65 units of deviation, while the 39 scores at 34 contribute 39(.35) = 13.65 units. The sum of these values divided by 60 gives the mean deviation.

$$\text{Mean Deviation} = \frac{13.65 + 13.65}{60} = \frac{27.3}{60} = .46$$

This figure, .46, indicates quantitatively the variability in this set of 60 scores. When compared with a set of data such as shown in Table 62, the mean deviation of which was found to be 7.35, the difference in spread is seen to be reflected in the size of the mean deviation figures. In interpreting the mean deviation, it will be helpful to think of it as defining two points on either side of the mean which enclose a large proportion of the scores (values). In a normal distribution, the mean deviation, when measured off on the scale above and below the mean, will mark the limits of the middle 57 per cent of the scores.

The Standard Deviation

The Standard Deviation is the most widely used measure of variability. It is defined as the square root of the mean square deviation. Defined by formula:

$$\text{Standard Deviation} = S.D. = \sigma = \sqrt{\frac{\Sigma D^2}{N}}$$

where ΣD^2 is read "the sum of the squared deviations of the scores from their mean" and N is the number of cases. S.D. and σ are abbreviations for the Standard Deviation. They are used interchangeably.

Although the Standard Deviation may be computed directly from a set of raw data by means of the formula $S.D. = \sqrt{\frac{\Sigma D^2}{N}}$, this process is laborious. For example, in the case of the set of data we have been using for illustrative purposes (tabulated in Table 55),

we would proceed by determining the difference between each raw score and the mean of the 60 scores, squaring these differences, summing the 60 squared differences, dividing by 60, and extracting the square root of the quotient. The first score tabulated is 15. The difference between this value and the mean of the 60 scores (as

TABLE 63

COMPUTATION OF THE STANDARD DEVIATION FROM A FREQUENCY DISTRIBUTION[3]

c.i.	f	d	fd	fd^2
51–53	2	6	12	72
48–50	2	5	10	50
45–47	3	4	12	48
42–44	4	3	12	36
39–41	6	2	12	24
36–38	9	1	9	9
33–35	9	0	0	0
30–32	7	−1	−7	7
27–29	5	−2	−10	20
24–26	4	−3	−12	36
21–23	4	−4	−16	64
18–20	2	−5	−10	50
15–17	2	−6	−12	72
12–14	1	−7	−7	49
			$\Sigma fd = -7$	$\Sigma fd^2 = 537$

$$\text{Mean} = M^\circ + \text{c.i.}\left(\frac{\Sigma fd}{N}\right)$$

$$\text{Mean} = 34 + 3\left(\frac{-7}{60}\right) = 33.65$$

$$\begin{aligned}\text{S.D.} &= \sqrt{\frac{\Sigma D^2}{N}} = \text{c.i.}\sqrt{\frac{\Sigma fd^2}{N} - \left(\frac{\Sigma fd}{N}\right)^2} \\ &= 3\sqrt{\tfrac{537}{60} - (.117)^2} \\ &= 3\sqrt{8.950 - .014} \\ &= 3\sqrt{8.936} \\ &= 3(2.99) \\ \text{S.D.} &= 8.97\end{aligned}$$

computed directly from the raw data) is $D = 33.6 - 15.0 = 18.6$. D^2 would therefore be $(18.6)^2 = 345.96$. This must be repeated for every one of the 60 scores before the sum of the squared deviations can be obtained.

---

[3] The derivation of the formula for computing the standard deviation by the method used in Table 63 is given below. Working in c.i. units rather than raw score units, let

$D_1, D_2 \ldots D_n$ = deviations of the scores in c.i. units from their mean.

c = difference in c.i. units between the mean of the scores and the midpoint of the zero c.i., i.e. the arbitrary base.

$d_1, d_2, \cdots d_n$ = deviations of the scores in c.i. units from the arbitrary base.

Then

$$\begin{aligned}D_1 &= d_1 - c \\ D_2 &= d_2 - c \\ &\cdots\cdots \\ D_n &= d_n - c\end{aligned}$$

If the above equations are squared on both sides, we have

$$\begin{aligned}D_1^2 &= (d_1 - c)^2 = d_1^2 - 2d_1c + c^2 \\ D_2^2 &= (d_2 - c)^2 = d_2^2 - 2d_2c + c^2 \\ &\cdots\cdots \\ D_n^2 &= (d_n - c)^2 = d_n^2 - 2d_nc + c^2\end{aligned}$$

Summating the above to determine the sum of the squared differences, we have

Because of the excessive labor in computing the S.D. directly from the raw data, a simple process which approximates the true value of the S.D. has been developed. This is used in the computations shown in Table 63, in which is computed the Standard Deviation of the data shown previously in Table 59.

$$D_1^2 + D_2^2 + \cdots D_n^2 = (d_1^2 + d_2^2 + \cdots d_n^2) + (2d_1c + 2d_2c + \cdots 2d_nc) + Nc^2$$

The above, written with summation signs, becomes:

$$\Sigma D^2 = \Sigma d^2 - 2c\Sigma d + Nc^2 \tag{1}$$

Now it will be remembered from page 491 that c, which is the difference in c.i. units between the mean of the data and the arbitrary base (i.e. the mean in c.i. units away from the arbitrary base) was determined by adding the d values of the scores and dividing this sum by the number of scores. Since there are f_1 scores at d_1 deviation; f_2 scores at d_2 deviation; etc.; this summation is given by $f_1d_1 + f_2d_2 + \cdots f_nd_n = \Sigma fd$, which, divided by N gives $\frac{\Sigma fd}{N}$.

We may therefore substitute in (1) $\frac{\Sigma fd}{N}$ for c, giving

$$\Sigma D^2 = \Sigma d^2 - 2\frac{\Sigma fd}{N}\Sigma d + N\left(\frac{\Sigma fd}{N}\right)^2$$

But the Σd and Σfd are the same, since Σfd is only a simpler method of determining the Σd that involves grouping together all scores of the same d, and multiplying this d by f, the number of such scores. For the same reason, Σd^2 is the same as Σfd^2. The above equation therefore may be written:

$$\Sigma D^2 = \Sigma fd^2 - 2\frac{\Sigma fd\Sigma fd}{N} + N\left(\frac{\Sigma fd}{N}\right)^2$$

$$\Sigma D^2 = \Sigma fd^2 - 2\frac{(\Sigma fd)^2}{N} + \frac{(\Sigma fd)^2}{N}$$

If both sides of the above equation are divided by N, we have

$$\frac{\Sigma D^2}{N} = \frac{\Sigma fd^2}{N} - 2\left(\frac{\Sigma fd}{N}\right)^2 + \left(\frac{\Sigma fd}{N}\right)^2$$

$$\frac{\Sigma D^2}{N} = \frac{\Sigma fd^2}{N} - \left(\frac{\Sigma fd}{N}\right)^2$$

Extracting the square root of both sides

$$S.D. = \sqrt{\frac{\Sigma D^2}{N}} = \sqrt{\frac{\Sigma fd^2}{N} - \left(\frac{\Sigma fd}{N}\right)^2}$$

which is the equation used in the computations accompanying Table 63 to determine the standard deviation of the distribution in c.i. units. The value yielded by this expression is then multiplied by the size of the c.i. to give the $S.D.$ in raw score units.

The standard deviation is the most commonly used measure of variability. Usually when the mean value of a set of data is given, the S.D. is also given to indicate the variability of the data.

Comparable scores

The S.D. performs another useful function—it can be used in the comparison of individual scores from different distributions. For example, suppose that two inspectors from departments A and B, who are working at different inspection jobs, detect respectively 45 and 89 defects during a week of work. How can we compare the efficiency of these two employees? It will be seen immediately that a direct comparison of the figures 45 and 89 is not valid, because the two inspection jobs may be very different. It will also be seen that we can say little concerning the position of these inspectors in their respective groups without knowing their relation to the mean of their group in inspection work. To make a comparison, then, we must first compute the mean number of defects spotted by all inspectors in Department A, and the mean number spotted by all inspectors in Department B. Suppose that those means are respectively 38 and 95. We thus see that the inspector from Department A is $45 - 38 = 7$ pieces *above* the mean for that department, and that the inspector from Department B is $89 - 95 = -6$, or 6 pieces *below* the mean of inspectors from that department. We can thus say, at this point, that the inspector from Department A is above average in ability on the job and that the inspector from Department B is below average. But how about their relative distance from the average? To answer this question we must compute the S.D.'s of the two distributions and determine how many S.D.'s each inspector is above or below average.

Suppose we find the S.D. of the operators in Department A to be 5.5 pieces. Our first inspector is therefore $\frac{45 - 38}{5.5} = 1.27$ S.D.'s above average. If the S.D. of the inspectors in Department B is 9.5, the inspector from the group who detected 89 pieces is $\frac{89 - 95}{9.5} = -.63$ or .63 S.D.'s below average. *The deviation of a score from the mean of the distribution expressed in S.D. units results in a measurement which is comparable with similarly determined measurements from other distributions.* Thus, we may say that our first inspector is about *twice* as far above average, in terms of comparable scale units, as the second operator is below average.

Scores computed in this manner are known as Z-scores. The formula for a Z-score is as follows:

$$\text{Z-score} = \frac{\text{Raw Score} - \text{Mean of Raw Scores}}{S.D. \text{ of Raw Scores}}$$

The Z-score is helpful not only when comparing scores from one distribution to another, but also when, for any reason, it is desired to combine scores with the same or differential weighting. A typical example of an industrial situation that requires this technique is in the combination of items used in a merit-rating blank. Suppose that each employee has been rated by his supervisor on a chart containing items such as the following.

	0–10	10–20	20–30	30–40	40–50
Industriousness	Always loafs if not watched	Often loafs when not watched	Sometimes loafs and sometimes works when not watched	Usually is hard at work	Always is hard at work
Knowledge of Job	Knows little about the job	Knows routine only	Is fairly well informed on his work	Well informed on details relative to his work	Thorough knowledge of present job and related work

We may suppose, for purposes of illustration, that it is now desired to combine these two traits into an overall merit rating. (If more than two traits are included in the chart, as is usually the case, the procedure is identical.) Suppose that an employee, Mr. A, has received 40 points on industriousness and 30 points on knowledge of job, making a total of 70 points if the ratings are added directly. Suppose that another employee, Mr. B, has received 30 points on industriousness and 40 points on knowledge of job, which also results in a total of 70 points if added directly. It is clear that such direct and immediate combination of ratings would result in identical overall ratings for these two employees. The question which we may raise is whether such a statement of equal ratings is justified. The answer is that it is not. If the mean rating of all employees on industriousness was 33 with a S.D. of 3, then *A*'s

rating would be $\frac{40 - 33}{3} = 2.33$, or 2.33 S.D.'s above the mean and B's would be $\frac{30 - 33}{3} = -1.00$, or 1.00 S.D. below the mean on this trait. If the mean rating for all employees on knowledge of job were 25, with a S.D. of 6, A would be $\frac{30 - 25}{6} = +.83$, or .83 S.D.'s above average in knowledge of job while B would be $\frac{40 - 25}{6}$ $= 2.50$, or 2.50 S.D.'s above average in this respect. Now, the proper combination of the two traits, if we wish to weight them equally, would be:

Employee	*Rating in Industriousness*	*Rating in Knowledge of Job*	*Z-Score in Industriousness*	*Z-Score in Knowledge of Job*	*Sum of Scores for Both Units*
A	40	30	+2.33	+.83	+3.16
B	30	40	−1.00	+2.50	+1.50

This transfer of ratings into Z-scores and the adding of the Z-scores shows that the two employees A and B are not equal in rating (as we would infer if the raw ratings were added), but rather that A is definitely higher than B. The procedure described has assumed that the two trait ratings, being combined, should be given equal weight, and the procedure shows how they can be combined with equal weight into a composite score. One might think that conversion of raw scores to Z-scores is not necessary if the raw scores are to be given equal weight in the combination score. Actually, if we do not give the raw scores equal weight in converting them into Z-scores, the scores will weight themselves according to the size of their respective standard deviations. In other words, if combined directly, the raw scores will be weighted too much or too little, depending upon their position relative to the means of their respective distributions and upon the variability of the distribution of which they are a part. When scores are combined, they are *always* weighted in some manner, whether we deliberately weight them or not. It is highly important, therefore, to weight them deliberately (either with equal weight or otherwise) by converting them into Z-scores and then combining them.

It does not follow from the above discussion that combined scores should always be weighted equally. Indeed, it is often desirable to weight various scores according to some plan that has been decided upon before the scores are combined. When this is

desired, such weighting can be accomplished very easily by multiplying each Z-score by the appropriate weight before they are combined. In our illustrative case, suppose that we have decided that *industriousness* should be given twice as much weight as *knowledge of job* in determining the total rating. This would be accomplished as follows:

Employee	Z-Score in Industriousness	Z-Score in Knowledge of Job	Weighted Z-Score in Industriousness	Weighted Z-Score in Knowledge of Job	Combined Weighted Z-Scores
A	+2.33	+.83	+4.66	+.83	+5.49
B	−1.00	+2.50	−2.00	+2.50	+.50

The combined ratings so obtained show a still greater difference between employees A and B than was obtained when the scores were equally weighted. If, on the other hand, it was desired to give the rating in knowledge of job twice as much weight as the rating on industriousness, the following computations would be made:

Employee	Z-Score in Industriousness	Z-Score in Knowledge of Job	Weighted Z-Score in Industriousness	Weighted Z-Score in Knowledge of Job	Combined Weighted Z-Score
A	+2.333	+.833	+2.333	+1.666	+4.0
B	−1.000	+2.500	−1.000	+5.000	+4.0

This last procedure results in giving identical total scores to employees A and B, which, it will be remembered, also occurred when the raw scores were added directly $(30 + 40 = 70)$. Now, if we remember that the assumed standard deviation for the knowledge of job ratings was 6, which is twice as great as the standard deviation of 3 assumed for the industriousness ratings, we can see why a direct combination of raw scores gives the equality that is obtained when the respective Z-scores are weighted in the ratio 2:1. In this instance the original ratings whose Standard Deviation is 6 are automatically given twice as much weight (when raw scores are combined directly) as the original ratings whose standard deviation

is only 3. Thus, we obtain the same final result (equality) by direct combination that is obtained when the *knowledge-of-job* ratings are deliberately given a weight twice as great as the industriousness ratings.

Many other problems arise in which it is necessary to weight scores to achieve a particular result. For example, a company faced with the problem of selecting a number of electrical apprentices desires to give this training to those boys who have the greatest aptitude for the job and who are therefore most likely to succeed. Careful consideration and discussion of the problem by management and supervision resulted in the decision that four factors should determine whether an employee should be given this training. These four factors were general intelligence, present knowledge of electricity, previous merit rating, and seniority with the company. It was further decided in conference that, although all of these four factors should be considered, they are not of equal importance. It was decided that a fair weighting of their relative importance was as follows:

General intelligence	40%
Knowledge of electricity	30%
Merit rating	20%
Seniority or service with the company	10%

To score the applicants according to this plan, each was given a general intelligence test and a test covering technical phases of electricity. Merit ratings and seniority were obtained from the company records. Each of the four scores was converted into a Z-score and the four resulting Z-scores were respectively multiplied by 40, 30, 20, and 10. For each employee the sum of the weighted Z-scores was used in indicating whether or not he was given the apprenticeship training.

Percentiles

The discussion of comparable scores should have made clear the fact that a raw score on any test is relatively meaningless unless it is interpreted in terms of its location in a distribution of other scores made by other people. If a test consists of 75 very easy questions, a score of 65 might be near the bottom of the distribution, and hence should be interpreted as a very low score. On the other hand, if a test consists of 75 very difficult questions, a score of 65 might be at or near the top of the distribution, and should therefore be considered a very high score. In other words, a raw score

of 65 might be a low score or a high score, depending upon the distribution of scores from which it is drawn.

One convenient and widely used method of interpreting a raw score is by using *percentile ranks*. A percentile rank may be defined as the number showing the percentage of the total group equal to or below the score in question. Thus, on a certain test, if 65 per cent of the total group scored 129 or below, the score of 129 would be at the 65th percentile, or would have a percentile rank of 65. The 50th percentile, it will be noted, is the same as the median as previously defined.

A convenient, practical method of determining by close approximation the percentile equivalents of a set of raw scores makes use of a cumulative frequency distribution such as the one tabulated in Table 64. This tabulation is based on the same distribution previously used in Table 57.

TABLE 64

CUMULATIVE FREQUENCY DISTRIBUTION USED IN DETERMINING PERCENTILE RANKS OF RAW SCORES

(1) *Class Intervals*	(2) *f*	(3) *Cumulative f*	(4) *Per Cent*
51–53	2	60	100
48–50	2	58	96
45–47	3	56	93
42–44	4	53	88
39–41	6	49	81
36–38	9	43	71
33–35	9	34	56
30–32	7	25	42
27–29	5	18	30
24–26	4	13	22
21–23	4	9	15
18–20	2	5	8
15–17	2	3	5
12–14	1	1	2
		Total = 60	

The per cent values, in column (4) of Table 64, are obtained by dividing each of the values in the cumulative *f* column, column (3), by the total of column (3), in this instance 60. The per cent values in column (4) are then plotted against the midpoints of the class intervals, as shown in Figure 137. From Figure 137 the percentile ranks of the raw scores may be read directly, with sufficient accuracy for most purposes. When this has been done, the raw scores with corresponding percentile ranks may be tabulated in a table, as illustrated in Table 65.

In Table 65, only certain percentile ranks are given, because these are usually sufficient for ordinary purposes. However, the intervening percentile ranks may also be read from Figure 137, if this is necessary.

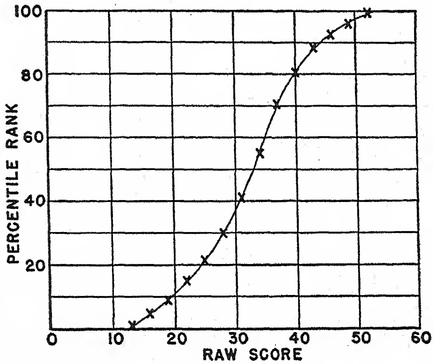


FIG. 137—Chart for converting raw scores into percentile ranks for illustrative data in Table 64.

The manual published with standardized tests usually includes percentile tables making possible the conversion of raw scores into percentile ranks. Since the percentile rank of a given raw score is

TABLE 65

RAW SCORES AND PERCENTILE RANK EQUIVALENTS READ FROM FIGURE 137

Percentile Rank	*Raw Score*
100	53
98	52
95	48
90	44
80	40
70	37
60	35
50	33
40	31
30	28
20	24
10	20
5	17
2	14
1	13

dependent upon the nature of the group used in constructing the conversion table, several raw-score-to-percentile-rank conversion tables, based on different groups, are often published with standardized tests.

Correlation

In numerous experimental situations, two variable quantities are so related that they vary, or tend to vary, with each other. A common problem in industrial psychology is to reduce to a simple and meaningful statement the facts that have been discovered concerning such a functional relationship. Suppose that a number of punch-press operators during a given period of time have each punched a certain number of pieces and have each mispunched, or otherwise wasted, a certain number of pounds of stock material. In such a situation, it might be of considerable importance for management to know whether any relationship exists (and, if so, how much) between quantity of work done and amount of material wasted. Indeed, the company's policy with respect to speed of work recommended as well as the quality control in the form of penalty or bonus might well be formulated correctly only in the light of specific knowledge of the relationship between speed and accuracy.

Consider a department employing eight operators for whom the following figures for production and waste in pounds are available:

Operator	*Production*	*Waste*
1.......	95	3.0
2.......	103	4.5
3.......	88	3.5
4.......	98	4.0
5.......	93	3.0
6.......	107	4.5
7.......	114	4.0
8.......	106	5.0

It is difficult, if not impossible, to determine from a gross inspection of these two columns of figures whether any relationship exists between speed and accuracy. It is necessary to employ some type of graphic or computational procedure to determine the amount of relationship which may exist between these two sets of data. One simple and sometimes satisfactory method consists of a simple plotting of the values on co-ordinate axes and rough inspection of the results. If we let production be represented on the X or horizontal axis, and waste on the Y or vertical axis, then the production and waste of each operator will locate him on a chart, giving the result shown in Figure 138.

A plot such as is shown in Figure 138 gives a much better indication of the presence or absence of a relationship between the data than can be obtained from the columns of raw data from which the chart was prepared. The chart shows that some relationship does

exist, and it is even possible to draw in by inspection a line or curve that represents this relationship in an approximate form.

Although this simple method of studying the relationship between two variables is sometimes adequate for very simple problems or for those that involve only a small amount of data, it is not adequate for an exact study because it does not result in a quantitative statement of the degree of relationship. The slope of the dotted line cannot be considered such a quantitative statement because: (1) this line is drawn in by inspection and, (2) its slope depends upon the units of measurement on both the X and Y axes.

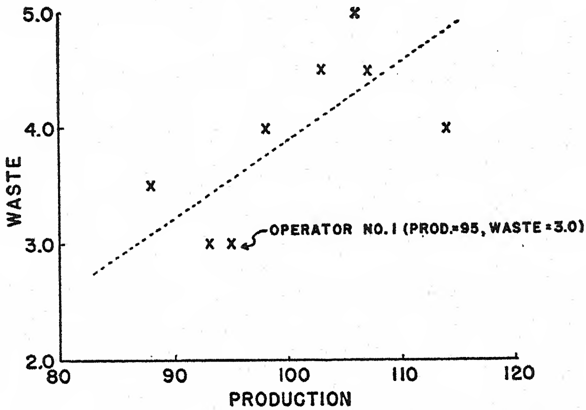


FIG. 138—A plot of the production and waste records for the eight punch-press operators shown above.

Two commonly used quantitative methods for measuring the degree of relationship between two paired sets of data are rank-order correlation and product-moment correlation.

Rank-order correlation

The use of this method may be described by applying it to the data for the eight punch press operators.

In Table 66 the two columns headed *Rank* give, respectively, the rank of the operators on the two measures (production and waste). The highest producing operator (in this case the seventh in the list) is given a rank of 1, the second highest a rank of 2, and so on. In like manner, the rank of each operator in wastage is placed in the waste-rank column. In case two or more operators are tied for a given rank (as in the case of the second and sixth operators who are tied at 4.5 pounds of waste each), the tied

scores are all given the same rank, which is the average of the ranks that would have been assigned to the tied scores if they had not been tied. The values in the D^2 column are obtained by squaring each D value. The sum of the D^2 column is then determined and the correlation computed by means of the formula:

$$R = 1 - \frac{6\Sigma D^2}{N(N^2 - 1)}$$

N is the number of cases entering into the computation.

This formula for the rank-order correlation is an empirical formula. It yields a value of +1.00 if the data are in exactly the same rank order. (The reason for this may be seen from the fact

TABLE 66

COMPUTATION OF RANK-ORDER CORRELATION RANK

Operator	*Production*	*Waste*	*Rank in Production*	*Rank in Waste*	*Difference in Rank* (D)	$(D)^2$
1.......	95	3.0	6	7.5	1.5	2.25
2.......	103	4.5	4	2.5	1.5	2.25
3.......	88	3.5	8	6.0	2.0	4.00
4.......	98	4.0	5	4.5	.5	.25
5.......	93	3.0	7	7.5	.5	.25
6.......	107	4.5	2	2.5	.5	.25
7.......	114	4.0	1	4.5	3.5	12.25
8.......	106	5.0	3	1.0	2.0	4.00
						25.50

$$R = 1 - \frac{6\Sigma D^2}{N(N^2 - 1)} = 1 - \frac{153}{504} = .70$$

that if all ranks are the same, all D's are zero, all D^2 values are zero, ΣD^2 is zero, and the formula becomes $1 - 0 = 1$.) If the data are in exactly reverse order (that is, if the individual who ranks highest on one series is lowest on the other, and so on) the formula will yield a value of −1.00, but if no relationship exists between the two sets of data, a correlation of zero will be found.

The use of this formula is ordinarily more satisfactory than a simple plotting of one variable against the other because it yields a quantitative statement of the degree of relationship and not simply a graphic representation that cannot be reduced to a numerical statement.

However, if an appreciable number of cases is involved, the rank-order method of computing the degree of relationship is extremely laborious. For this reason—and for other reasons of a mathematical nature—it is ordinarily used only when the data are limited to a very few cases (less than 30).

The product-moment coefficient

This is the most widely used measure of relationship. Like the rank-order correlation, it may vary from +1.00 (indicating perfect positive relationship) through zero (indicating no relationship) to −1.00 (indicating perfect negative relationship). The product-moment correlation, represented by the symbol r, may be defined in several ways. One of the simplest definitions is that *r is the slope of the straight line which best fits the data after the data have been plotted as*

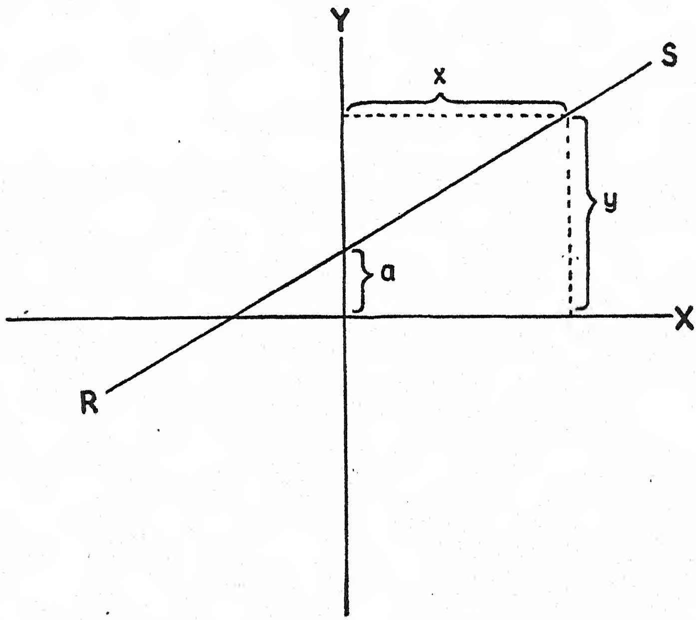


FIG. 139—The slope of the line $R\ S$ is defined as $b = \frac{y - a}{x}$.

z-scores on co-ordinate axes; that is, it is the tangent of the angle made by this line with the base line.

Several terms in this definition require further definition. By *slope* is meant steepness with which the line rises. The slope of a straight line drawn in any manner across co-ordinate paper is defined as the distance, y, from any given point on the line to the x intercept, minus the distance, a, from the origin to the y intercept, divided by the distance, x, from the point on the line to y intercept.

Thus the slope, which we will call b, is defined in Figure 139 as follows:

$$b = \frac{y - a}{x}$$

It should be remembered that, on co-ordinate axes, distances above and to the right of the origin are positive, while distances measured below and/or to the left of the origin are negative. The slope of any line which *rises* as it goes from left to right will therefore be positive (the greater the rise in a given distance to the right the larger the positive value of the slope) and the slope of any line which *falls* as it goes from left to right will be negative (the greater the fall in a given distance to the right, the greater the negative value of the slope).

By line of *best fit* in the definition is meant a line so drawn that the sum of the squared deviations in a vertical direction from the original points to the line is less than the sum would be for any other straight line that might be drawn.

A rough approximation of the value of r may be obtained by plotting the z-scores of the two variables, fitting a straight line to these points by inspection, and graphically measuring the slope of this straight line. Although this method is never used in practical computation (because it is both inaccurate and laborious), the application of it to a set of representative data may serve to clarify the meaning of the correlation coefficient, r. Returning to the data for which we have previously computed the rank-order correlation (see Table 66 on page 510), we first compute the z-scores for each measure:

TABLE 67

PRODUCTION AND WASTE FOR EIGHT PUNCH PRESS OPERATORS, WITH CORRESPONDING Z-SCORES OF THE PRODUCTION AND WASTE FIGURES

Operator	*Production*	*Waste*	*Z-Score in Production*	*Z-Score in Waste*
1	95	3.0	−.69	−1.38
2	103	4.5	+.31	+.82
3	88	3.5	−1.56	−.65
4	98	4.0	−.31	+.09
5	93	3.0	−.94	−1.38
6	107	4.5	+.81	+.82
7	114	4.0	+1.69	+.09
8	106	5.0	+.69	+1.56
Mean	100.5	3.94		
S.D	8.0	.68		

These pairs of z-scores are used as the x and y values for eight points which are plotted on co-ordinate axes as in Figure 140. The

straight line that seems best to fit these points is then determined (as with a stretched string which is moved about until the desired location is obtained) and drawn on the graph. The correlation, r, as determined by this crude method, is obtained by measuring the slope of this line. The procedure applied to Figure 140 gives a

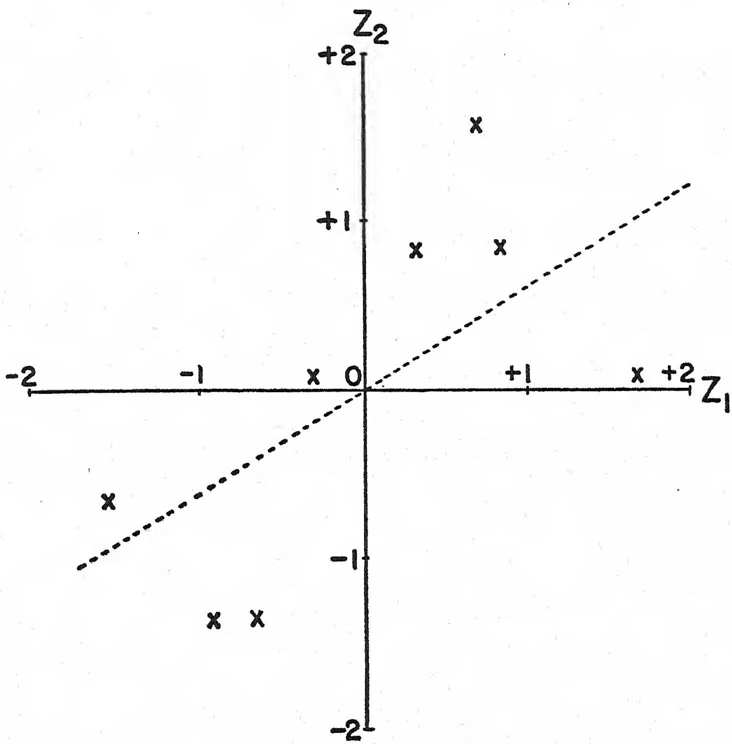


FIG. 140—A plot of the z-scores for production and waste records of the eight punch-press operators shown above.

value of $r = .61$, but it should be emphasized that this value is affected by:

1. The accuracy with which the straight line has been located, and

2. The accuracy with which the slope of the line has been measured after it has been drawn.

Both points (1) and (2) operate to eliminate the possibility of complete accuracy in this method of determining a correlation coefficient. Therefore, a mathematical method has been devised to

make the computation, so that no plotting of points or graphic measurements are required. This method involves determining the equation of the straight line which, if plotted, would best fit the points and computing the slope of this straight line of best fit from the equation.

TABLE 68

COMPUTATION OF r BY Z-SCORE METHOD BETWEEN PRODUCTION AND WASTE FIGURES

Operator	*Production*	*Waste*	*Z-Score in Production*	*Z-Score in Waste*	(Z_1Z_2)
1.......	95	3.0	−.69	−1.38	+.95
2.......	103	4.5	+.31	+.82	+.25
3.......	88	3.5	−1.56	−.65	+1.01
4.......	98	4.0	−.31	+.09	−.03
5.......	93	3.0	−.94	−1.38	+1.30
6.......	107	4.5	+.81	+.82	+.66
7.......	114	4.0	+1.69	+.09	+.15
8.......	106	5.0	+.69	+1.56	+1.08
					+5.38

$$r = \frac{Z_xZ_y}{N} = \frac{+5.38}{8} = .67$$

It may be proved mathematically that the slope of the straight line of best fit is given by the following equation: Slope $= r = \frac{\Sigma Z_xZ_y}{N}$,[4] where ΣZ_xZ_y is read "the sum of the products of the z-scores for the pairs of points or values."

---

[4] The proof of this formula is as follows:

$Z_{x1}, Z_{x2}, Z_{x3}, \cdots Z_{xn}$ are the z-scores for the x variable
$Z_{y1}, Z_{y2}, Z_{y3}, \cdots Z_{yn}$ are the z-scores for the y variable

The equation of any straight line is: $y = a + bx$. The problem is to find the values of the constants a and b in this equation which will give the straight line that "best fits" the data according to the criterion of best fit stated on page 512, that is, the straight line which will give a minimum value to the sum of the squared deviations between the line and the original points.

The first point (whose co-ordinate points are Z_{x1} and Z_{y1}) will deviate from the line by an amount:

$$d_1 = Z_{y1} - a - bZ_{x1}$$

this deviation squared will be

$$d_1^2 = (Z_{y1} - a - bZ_{x1})^2$$

If the sum of all such squared deviations of the points from the line be represented by u, then

$$u = \Sigma d^2 = \Sigma(Z_y - a - bZ_x)^2$$

For the values of a and b which result in u being a minimum, the derivatives of u

Applying the formula to the data in Table 66, we may compute the correlation as in Table 68.

The value of r thus obtained by computation, .67, differs from the value of .61 obtained by plotting and inspection. The plotting and inspection method yielded a value which was somewhat in error for the data in question.

---

with respect to a and b, respectively, must be zero. To find the correlation coefficient it is therefore only necessary to differentiate the equation with respect to a and to b, to set the resulting derivatives equal to zero, and to solve for b (which is the slope of the straight line of best fit). This is done as follows:

$$\frac{\partial u}{\partial a} = 0 = 2\Sigma(Z_y - a - bZ_x)(-1) \tag{1}$$

$$\frac{\partial u}{\partial b} = 0 = 2\Sigma(Z_y - a - bZ_x)(-Z_x) \tag{2}$$

(1) above becomes:

$$0 = -\Sigma Z_y + Na + b\Sigma Z_x$$
$$\Sigma Z_y = Na + b\Sigma Z_x \tag{3}$$

(2) above becomes:

$$0 = -\Sigma Z_x Z_y + a\Sigma Z_x + b\Sigma Z_x^2$$
$$\Sigma Z_x Z_y = a\Sigma Z_x + b\Sigma Z_x^2 \tag{4}$$

It will be remembered that a z-score is obtained as follows (see page 502):

$$Z_x = \frac{X - M_x}{\sigma x}$$

Where X is a given raw score, M_x is the mean raw score, and σ_x the standard deviation of the raw scores.
The sum of all z-scores is therefore:

$$\begin{aligned} \Sigma Z_x &= \frac{\Sigma(X - M_x)}{\sigma_x} \\ &= \frac{\Sigma X}{\sigma_x} - \frac{NM_x}{\sigma_x} \\ &= \frac{\Sigma X}{\sigma_x} - \frac{\frac{N\Sigma X}{N}}{\sigma_x} \\ &= \frac{\Sigma X}{\sigma_x} - \frac{\Sigma X}{\sigma_x} \\ &= 0 \end{aligned}$$

In a similar way it can be shown that

$$\Sigma Z_y = 0$$

By substituting 0 for ΣZ_x and ΣZ_y in (3) we find immediately that a in the equa-

While the z-score method of computing a correlation coefficient illustrated in Table 68 may be used with any number of cases and will yield the correct mathematical value of r, the use of this method

---

tion of the straight line of best fit is zero.

Working with (4), we find the value of b (which is the slope or the correlation coefficient) as follows:

$$\Sigma Z_x Z_y = a\Sigma Z_x + b\Sigma Z_x^2$$

since $\Sigma Z_x = 0$, this becomes:

$$\Sigma Z_x Z_y = b\Sigma Z_x^2$$

$$b = \frac{\Sigma Z_x Z_y}{\Sigma Z_x^2} \tag{5}$$

It may be shown as follows that $\Sigma Z_x^2 = N$

$$Z_x = \frac{X - M_x}{\sigma_x}$$

$$Z_x^2 = \frac{(X - M_x)^2}{\sigma_x^2}$$

$$\Sigma Z_x^2 = \frac{\Sigma(X - M_x)^2}{\sigma_x^2}$$

$$\Sigma Z_x^2 = \frac{\Sigma(X^2 - 2XM_x + M_x^2)}{\sigma_x^2}$$

$$\Sigma Z_x^2 = \frac{\Sigma X^2 - 2M_x\Sigma X + NM_x^2}{\sigma_x^2}$$

$$\Sigma Z_x^2 = \frac{\Sigma X^2 - 2\frac{\Sigma X}{N}\Sigma X + N\left(\frac{\Sigma X}{N}\right)^2}{\sigma_x^2}$$

$$\Sigma Z_x^2 = \frac{\frac{N\Sigma X^2 - 2(\Sigma X)^2 + (\Sigma X)^2}{N}}{\frac{\Sigma X^2}{N} - \left(\frac{\Sigma X}{N}\right)^2}$$

$$\Sigma Z_x^2 = \frac{\frac{N\Sigma X^2 - (\Sigma X)^2}{N}}{\frac{N\Sigma X^2 - (\Sigma X)^2}{N^2}}$$

$$\Sigma Z_x^2 = \frac{N^2(\Sigma X^2) - N(\Sigma X)^2}{N\Sigma X^2 - (\Sigma X)^2}$$

$$\Sigma Z_x^2 = \frac{N[N\Sigma X^2 - (\Sigma X)^2]}{[N\Sigma X^2 - (\Sigma X)^2]}$$

$$\Sigma Z_x^2 = N \tag{6}$$

Substituting the value of $\Sigma(Z_x)^2$ given in (6) in equation (5), we have the slope or $r = b = \frac{\Sigma Z_x Z_y}{N}$.

when many pairs of data are to be correlated is very laborious. It is therefore recommended, under such circumstances, that a modification of the fundamental formula $r = \frac{\Sigma Z_x Z_y}{N}$ which makes it possible to compute r from raw score values rather than z-score values be used. One convenient formula[5] for determining the coefficient of correlation directly from the raw data is:

$$r = \frac{N\Sigma XY - \Sigma X \Sigma Y}{\sqrt{N\Sigma X^2 - (\Sigma X)^2}\sqrt{N\Sigma Y^2 - (\Sigma Y)^2}}$$

When we apply this formula, for illustrative purposes, to the

---

[5] The proof of this formula is as follows:

$$r = \frac{\Sigma Z_x Z_y}{N}$$

$$= \frac{1}{N}\Sigma \frac{(X - M_x)}{\sigma_x}\frac{(Y - M_y)}{\sigma_y}$$

$$= \frac{1}{N}\Sigma \frac{(XY - XM_y - YM_x + M_xM_y)}{\sigma_x\ \sigma_y}$$

$$= \frac{1}{N}\Sigma \frac{\left(XY - \frac{X\Sigma Y}{N} - Y\frac{\Sigma X}{N} + \frac{\Sigma X \Sigma Y}{N^2}\right)}{\sigma_x\ \sigma_y}$$

$$= \frac{1}{N}\frac{\Sigma XY - \frac{\Sigma X \Sigma Y}{N} - \frac{\Sigma Y \Sigma X}{N} + \frac{N\Sigma X \Sigma Y}{N^2}}{\sqrt{\frac{\Sigma X^2}{N} - \left(\frac{\Sigma X}{N}\right)^2}\sqrt{\frac{\Sigma Y^2}{N} - \left(\frac{\Sigma Y}{N}\right)^2}}$$

$$= \frac{1}{N}\frac{\frac{N\Sigma XY - \Sigma X \Sigma Y}{N}}{\sqrt{\frac{N\Sigma X^2 - (\Sigma X)^2}{N^2}}\sqrt{\frac{N\Sigma Y^2 - (\Sigma Y)^2}{N^2}}}$$

$$= \frac{1}{N}\frac{\frac{N\Sigma XY - \Sigma X \Sigma Y}{N}}{\frac{\sqrt{N\Sigma X^2 - (\Sigma X)^2}}{N}\frac{\sqrt{N\Sigma Y^2 - (\Sigma Y)^2}}{N}}$$

$$= \frac{1}{N}\frac{N^2(N\Sigma XY - \Sigma X \Sigma Y)}{N\sqrt{N\Sigma X^2 - (\Sigma X)^2}\sqrt{N\Sigma Y^2 - (\Sigma Y)^2}}$$

$$= \frac{N\Sigma XY - \Sigma X \Sigma Y}{\sqrt{N\Sigma X^2 - (\Sigma X)^2}\sqrt{N\Sigma Y^2 - (\Sigma Y)^2}}$$

data tabulated in Table 68, the computations shown in Table 69 result:

TABLE 69

COMPUTATION OF r DIRECTLY FROM RAW DATA

Operator	*Production* (x)	*Waste* (y)	X^2	Y^2	XY
1.....	95	3.0	9025	9.00	285.0
2.....	103	4.5	10609	20.25	463.5
3.....	88	3.5	7744	12.25	308.0
4.....	98	4.0	9604	16.00	392.0
5.....	93	3.0	8649	9.00	279.0
6.....	107	4.5	11449	20.25	481.5
7.....	114	4.0	12996	16.00	456.0
8.....	106	5.0	11236	25 00	530.0
	$\Sigma X = 804$	$\Sigma Y = 31.5$	$\Sigma X^2 = 81312$	$\Sigma Y^2 = 127.75$	$\Sigma XY = 3195.0$

$$r = \frac{N\Sigma XY - \Sigma X \Sigma Y}{\sqrt{N\Sigma X^2 - (\Sigma X)^2}\sqrt{N\Sigma Y^2 - (\Sigma Y)^2}}$$

$$r = \frac{8(3195) - (804)(31.5)}{\sqrt{8(81312) - (804)^2}\sqrt{8(127.75) - (31.5)^2}}$$

$$r = .67$$

When a considerable number of pairs of data are to be correlated, the use of a chart will still further simplify the computations. Several forms of such a chart have been prepared. One convenient form is shown in Figure 141. This chart shows the computation of the correlation between time used in inspecting 300 pieces of material and the number of defective pieces detected. In using this chart the following steps should be followed:

1. Decide upon appropriate class intervals for one of the variables (using the rules given on page 486) and write these in on either the x or the y axis.
2. Decide upon appropriate class intervals for the other variables and write these in on the axis not used in (1) above.
3. Place one tally mark on the scattergram for each pair of values being correlated. For example, if an inspector spotted 33 defects in 16.5 minutes, the tally mark would go in the pigeonhole that is found at the intersection of the *row* containing 33 defects and the *column* containing 16.5 minutes.
4. After all tally marks have been placed on the chart, the rows should be added horizontally and the sum of the tally marks in each row written opposite this row in the f_y column (Column *1*).
5. The tally marks in each column should be added and the sum written at the bottom of each column in the row (A), the f_x row.
6. The f_y column should be added and the sum written oppo-

site N at the bottom of this column. *The value of N thus obtained may be checked by adding the values in the f_x row. The sum of these values should also give the value of N.*

7. Each value of f_y in the column so headed should be multiplied by the value of d_y opposite it, and the resultant product written

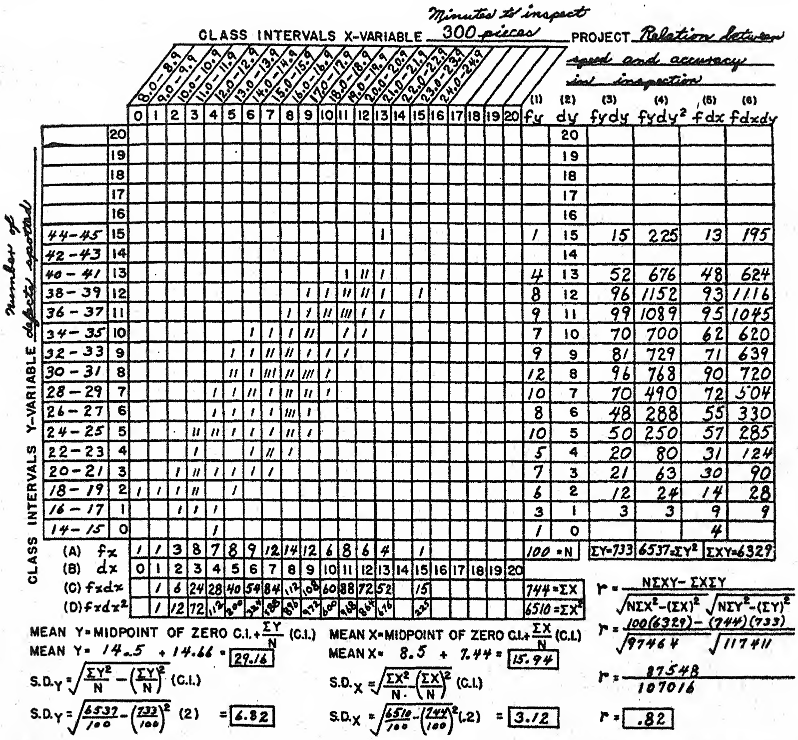


FIG. 141—A chart used in the computation of a product moment coefficient of correlation.

in Column *3*, headed $f_y d_y$. *The sum of Column 3 is the value of ΣY, which is used in the formula.*

8. Each value in Column *3*, the $f_y d_y$ column, should be multiplied by the corresponding value in Column *2*, the d_y column, resulting in the values for Column *4*, or the $f_y d_y^2$ column. *The sum of Column 4 is the value of ΣY^2 which is used in the formula.*

9. The values going into Column *5*, the fd_x column, are determined by finding, for each row, the sum of the products of the number of cases in each cell times the x value of that cell. For

example, in the first row in which a tally mark appears, there is only a single case, which appears in the cell under an x value of 13. The value to go into the blank in Column *5* is therefore (1) (13) = 13. In the next row no tally marks appear; therefore this row is blank. In the next row, one tally mark appears in the cell under an x value of 11, two in the cell with an x value of 12, and one in the cell with an x value of 13. The value to go into the blank cell in Column *5* is therefore (1) (11) + (2) (12) + (1) (13) = 48. The remaining cells in Column *5* are filled in a similar manner.

10. The cells in Column *6*, the fd_xy_y column, are filled with values obtained by multiplying each value in Column *2*, the d_y column, by the value in that same row appearing in Column *5*, the fd_x column. The value in the first cell in Column *6* is therefore (15) (13) = 195. *The sum of Column 6 is the value of* ΣXY *which is used in the formula.*

11. The values in row (C) are obtained by multiplying each value in row (A), the f_x row, by the value directly below it in row (B), the d_x row. The values appearing in row A have already been obtained (see Step *5* above). The resultant values are entered in row (C), the f_xd_x row. *The sum of the values appearing in row* (C) *is the value of* ΣX *which is used in the formula.*

12. Each value in row (B), the d_x row, should be multiplied by the value directly below in row (C), the f_xd_x row. The resultant values should be entered in row (D), the $f_xd_x^2$ row. *The sum of the values in row* (D) *is the value of* ΣX^2 *which is used in the formula.*

13. The values for N (see Step *6*), ΣY (see Step *7*), ΣY^2 (see Step *8*), ΣXY (see Step *10*), ΣX (see Step *11*), and ΣX^2 (see Step *12*) are now entered in the formula. The indicated arithmetic computations are then performed, yielding the value of r.

The use of this method assumes that each measure has the value of the midpoint of the class interval in which it falls. The computations indicated on the chart result in obtaining not only the value for r but also the mean and the standard deviation of both the X and Y arrays. It will be noted that these are the same formulas previously considered on pages 499 and 500.

Appendix B

Taylor-Russell Tables[1]

Tables of the Proportion Who Will Be Satisfactory Among Those Selected, for Given Values of the Proportion of Present Employees Considered Satisfactory, the Selection Ratio, and r

Proportion of Employees Considered Satisfactory = .05

Selection Ratio

r	.05	.10	.20	.30	.40	.50	.60	.70	.80	.90	.95
.00	.05	.05	.05	.05	.05	.05	.05	.05	.05	.05	.05
.05	.06	.06	.06	.06	.06	.05	.05	.05	.05	.05	.05
.10	.07	.07	.07	.06	.06	.06	.06	.05	.05	.05	.05
.15	.09	.08	.07	.07	.07	.06	.06	.06	.05	.05	.05
.20	.11	.09	.08	.08	.07	.07	.06	.06	.06	.05	.05
.25	.12	.11	.09	.08	.08	.07	.07	.06	.06	.05	.05
.30	.14	.12	.10	.09	.08	.07	.07	.06	.06	.05	.05
.35	.17	.14	.11	.10	.09	.08	.07	.06	.06	.05	.05
.40	.19	.16	.12	.10	.09	.08	.07	.07	.06	.05	.05
.45	.22	.17	.13	.11	.10	.08	.08	.07	.06	.06	.05
.50	.24	.19	.15	.12	.10	.09	.08	.07	.06	.06	.05
.55	.28	.22	.16	.13	.11	.09	.08	.07	.06	.06	.05
.60	.31	.24	.17	.13	.11	.09	.08	.07	.06	.06	.05
.65	.35	.26	.18	.14	.11	.10	.08	.07	.06	.06	.05
.70	.39	.29	.20	.15	.12	.10	.08	.07	.06	.06	.05
.75	.44	.32	.21	.15	.12	.10	.08	.07	.06	.06	.05
.80	.50	.35	.22	.16	.12	.10	.08	.07	.06	.06	.05
.85	.56	.39	.23	.16	.12	.10	.08	.07	.06	.06	.05
.90	.64	.43	.24	.17	.13	.10	.08	.07	.06	.06	.05
.95	.73	.47	.25	.17	.13	.10	.08	.07	.06	.06	.05
1.00	1.00	.50	.25	.17	.13	.10	.08	.07	.06	.06	.05

Proportion of Employees Considered Satisfactory = .10

Selection Ratio

r	.05	.10	.20	.30	.40	.50	.60	.70	.80	.90	.95
.00	.10	.10	.10	.10	.10	.10	.10	.10	.10	.10	.10
.05	.12	.12	.11	.11	.11	.11	.11	.10	.10	.10	.10
.10	.14	.13	.13	.12	.12	.11	.11	.11	.11	.10	.10
.15	.16	.15	.14	.13	.13	.12	.12	.11	.11	.10	.10
.20	.19	.17	.15	.14	.14	.13	.12	.12	.11	.11	.10
.25	.22	.19	.17	.16	.14	.13	.13	.12	.11	.11	.10
.30	.25	.22	.19	.17	.15	.14	.13	.12	.12	.11	.10
.35	.28	.24	.20	.18	.16	.15	.14	.13	.12	.11	.10
.40	.31	.27	.22	.19	.17	.16	.14	.13	.12	.11	.10
.45	.35	.29	.24	.20	.18	.16	.15	.13	.12	.11	.10
.50	.39	.32	.26	.22	.19	.17	.15	.13	.12	.11	.11
.55	.43	.36	.28	.23	.20	.17	.15	.14	.12	.11	.11
.60	.48	.39	.30	.25	.21	.18	.16	.14	.12	.11	.11
.65	.53	.43	.32	.26	.22	.18	.16	.14	.12	.11	.11
.70	.58	47	.35	.27	.22	.19	.16	.14	.12	.11	.11
.75	.64	.51	.37	.29	.23	.19	.16	.14	.12	.11	.11
.80	.71	.56	.40	.30	.24	.20	.17	.14	.12	.11	.11
.85	.78	.62	.43	.31	.25	.20	.17	.14	.12	.11	.11
.90	.86	.69	.46	.33	.25	.20	.17	.14	.12	.11	.11
.95	.95	.78	.49	.33	.25	.20	.17	.14	.12	.11	.11
1.00	1.00	1.00	.50	.33	.25	.20	.17	.14	.13	.11	.11

[1] These tables are reproduced by permission from H. C. Taylor and J. T. Russell, "The Relationship of Validity Coefficients to the Practical Effectiveness of Tests in Selection: Discussion and Tables," *Journal of Applied Psychology*, XXIII (1939), pp. 565–578.

Proportion of Employees Considered Satisfactory = .20
Selection Ratio

r	.05	.10	.20	.30	.40	.50	.60	.70	.80	.90	.95
00	.20	.20	.20	.20	.20	.20	.20	.20	.20	.20	.20
.05	.23	.23	.22	.22	.21	.21	.21	.21	.20	.20	.20
.10	.26	.25	.24	.23	.23	.22	.22	.21	.21	.21	.20
.15	.30	.28	.26	.25	.24	.23	.23	.22	.21	.21	.20
.20	.33	.31	.28	.27	.26	.25	.24	.23	.22	.21	.21
.25	.37	.34	.31	.29	.27	.26	.24	.23	.22	.21	.21
.30	.41	.37	.33	.30	.28	.27	.25	.24	.23	.21	.21
.35	.45	.41	.36	.32	.30	.28	.26	.24	.23	.22	.21
.40	.49	.44	.38	.34	.31	.29	.27	.25	.23	.22	.21
.45	.54	.48	.41	.36	.33	.30	.28	.26	.24	.22	.21
.50	.59	.52	.44	.38	.35	.31	.29	.26	.24	.22	.21
.55	.63	.56	.47	.41	.36	.32	.29	.27	.24	.22	.21
.60	.68	.60	.50	.43	.38	.34	.30	.27	.24	.22	.21
.65	.73	.64	.53	.45	.39	.35	.31	.27	.25	.22	.21
.70	.79	.69	.56	.48	.41	.36	.31	.28	.25	.22	.21
.75	.84	.74	.60	.50	.43	.37	.32	.28	.25	.22	.21
.80	.89	.79	.64	.53	.45	.38	.33	.28	.25	.22	.21
.85	.94	.85	.69	.56	.47	.39	.33	.28	.25	.22	.21
.90	.98	.91	.75	.60	.48	.40	.33	.29	.25	.22	.21
.95	1.00	.97	.82	.64	.50	.40	.33	.29	.25	.22	.21
1.00	1.00	1.00	1.00	.67	.50	.40	.33	.29	.25	.22	.21

Proportion of Employees Considered Satisfactory = .30
Selection Ratio

r	.05	.10	.20	.30	.40	.50	.60	.70	.80	.90	.95
.00	.30	.30	.30	.30	.30	.30	.30	.30	.30	.30	.30
.05	.34	.33	.33	.32	.32	.31	.31	.31	.31	.30	.30
.10	.38	.36	.35	.34	.33	.33	.32	.32	.31	.31	.30
.15	.42	.40	.38	.36	.35	.34	.33	.33	.32	.31	.31
.20	.46	.43	.40	.38	.37	.36	.34	.33	.32	.31	.31
.25	.50	.47	.43	.41	.39	.37	.36	.34	.33	.32	.31
.30	.54	.50	.46	.43	.40	.38	.37	.35	.33	.32	.31
.35	.58	.54	.49	.45	.42	.40	.38	.36	.34	.32	.31
.40	.63	.58	.51	.47	.44	.41	.39	.37	.34	32	.31
.45	.67	.61	.55	.50	.46	.43	.40	.37	.35	.32	.31
.50	.72	.65	.58	.52	.48	.44	.41	.38	.35	.33	.31
.55	.76	.69	.61	.55	.50	.46	.42	.39	.36	.33	.31
.60	.81	.74	.64	.58	.52	.47	.43	.40	.36	.33	.31
.65	.85	.78	.68	.60	.54	.49	.44	.40	.37	.33	.32
.70	.89	.82	.72	.63	.57	.51	.46	.41	.37	.33	.32
.75	.93	.86	.76	.67	.59	.52	.47	.42	.37	.33	.32
.80	.96	.90	.80	.70	.62	.54	.48	.42	.37	.33	.32
.85	.99	.94	.85	.74	.65	.56	.49	.43	.37	.33	.32
.90	1.00	.98	90	.79	.68	.58	.49	.43	.37	.33	.32
.95	1 00	1.00	.96	.85	.72	.60	.50	.43	.37	.33	.32
1.00	1.00	1.00	1.00	1 00	.75	.60	.50	.43	38	.33	.32

Proportion of Employees Considered Satisfactory = .40
Selection Ratio

r	.05	.10	.20	.30	.40	.50	.60	.70	.80	.90	.95
.00	.40	.40	.40	.40	.40	.40	.40	.40	.40	.40	.40
.05	.44	.43	.43	.42	.42	.42	.41	.41	.41	.40	.40
.10	.48	.47	.46	.45	.44	.43	.42	.42	.41	.41	.40
.15	.52	.50	.48	.47	.46	.45	.44	.43	.42	.41	.41
.20	.57	.54	.51	.49	.48	.46	.45	.44	.43	.41	.41
.25	.61	.58	.54	.51	.49	.48	.46	.45	.43	.42	.41
.30	.65	.61	.57	.54	.51	.49	.47	.46	.44	.42	.41
.35	.69	.65	.60	.56	.53	.51	.49	.47	.45	.42	.41
.40	.73	.69	.63	.59	.56	.53	.50	.48	.45	.43	.41
.45	.77	.72	.66	.61	.58	.54	.51	.49	.46	.43	.42
.50	.81	.76	.69	.64	.60	.56	.53	.49	.46	.43	.42
.55	.85	.79	.72	.67	.62	.58	.54	.50	.47	.44	.42
.60	.89	.83	.75	.69	.64	.60	.55	.51	.48	.44	.42
.65	.92	.87	.79	.72	.67	.62	.57	.52	.48	.44	.42
.70	.95	.90	.82	.76	.69	.64	.58	.53	.49	.44	.42
.75	.97	.93	.86	.79	.72	.66	.60	.54	.49	.44	.42
.80	.99	.96	.89	.82	.75	.68	.61	.55	.49	.44	.42
.85	1.00	.98	.93	.86	.79	.71	.63	.56	.50	.44	.42
.90	1.00	1.00	.97	.91	.82	.74	.65	.57	.50	.44	.42
.95	1.00	1.00	.99	.96	.87	.77	.66	.57	.50	.44	42
1.00	1.00	1.00	1.00	1.00	1.00	.80	.67	.57	.50	.44	.42

Proportion of Employees Considered Satisfactory = .50
Selection Ratio

r	.05	.10	.20	.30	.40	.50	.60	.70	.80	.90	.95
.00	.50	.50	.50	.50	.50	.50	.50	.50	.50	.50	.50
.05	.54	.54	.53	.52	.52	.52	.51	.51	.51	.50	.50
.10	.58	.57	.56	.55	.54	.53	.53	.52	.51	.51	.50
.15	.63	.61	.58	.57	.56	.55	.54	.53	.52	.51	.51
.20	.67	.64	.61	.59	.58	.56	.55	.54	.53	.52	.51
.25	.70	.67	.64	.62	.60	.58	.56	.55	.54	.52	.51
.30	.74	.71	.67	.64	.62	.60	.58	.56	.54	.52	.51
.35	.78	.74	.70	.66	.64	.61	.59	.57	.55	.53	.51
.40	.82	.78	.73	.69	.66	.63	.61	.58	.56	.53	.52
.45	.85	.81	.75	.71	.68	.65	.62	.59	.56	.53	.52
.50	.88	.84	.78	.74	.70	.67	.63	.60	.57	.54	.52
.55	.91	.87	.81	.76	.72	.69	.65	.61	.58	.54	.52
.60	.94	.90	.84	.79	.75	.70	.66	.62	.59	.54	.52
.65	.96	.92	.87	.82	.77	.73	.68	.64	.59	.55	.52
.70	.98	.95	.90	.85	.80	.75	.70	.65	.60	.55	.53
.75	.99	.97	.92	.87	.82	.77	.72	.66	.61	.55	.53
.80	1.00	.99	.95	.90	.85	.80	.73	.67	.61	.55	.53
.85	1.00	.99	.97	.94	.88	.82	.76	.69	.62	.55	.53
.90	1.00	1.00	.99	.97	.92	.86	.78	.70	.62	.56	.53
.95	1.00	1.00	1.00	.99	.96	.90	.81	.71	.63	.56	.53
1.00	1.00	1.00	1.00	1.00	1.00	1.00	.83	.71	.63	.56	.53

Proportion of Employees Considered Satisfactory = .60
Selection Ratio

r	.05	.10	.20	.30	.40	.50	.60	.70	.80	.90	.95
.00	.60	.60	.60	.60	.60	.60	.60	.60	.60	.60	.60
.05	.64	.63	.63	.62	.62	.62	.61	.61	.61	.60	.60
.10	.68	.67	.65	.64	.64	.63	.63	.62	.61	.61	.60
.15	.71	.70	.68	.67	.66	.65	.64	.63	.62	.61	.61
.20	.75	.73	.71	.69	.67	.66	.65	.64	.63	.62	.61
.25	.78	.76	.73	.71	.69	.68	.66	.65	.63	.62	.61
.30	.82	.79	.76	.73	.71	.69	.68	.66	.64	.62	.61
.35	.85	.82	.78	.75	.73	.71	.69	.67	.65	.63	.62
.40	.88	.85	.81	.78	.75	.73	.70	.68	.66	.63	.62
.45	.90	.87	.83	.80	.77	.74	.72	.69	.66	.64	.62
.50	.93	.90	.86	.82	.79	.76	.73	.70	.67	.64	.62
.55	.95	.92	.88	.84	.81	.78	.75	.71	.68	.64	.62
.60	.96	.94	.90	.87	.83	.80	.76	.73	.69	.65	.63
.65	.98	.96	.92	.89	.85	.82	.78	.74	.70	.65	.63
.70	.99	.97	.94	.91	.87	.84	.80	.75	.71	.66	.63
.75	.99	.99	.96	.93	.90	.86	.81	.77	.71	.66	.63
.80	1.00	.99	.98	.95	.92	.88	.83	.78	.72	.66	.63
.85	1.00	1.00	.99	.97	.95	.91	.86	.80	.73	.66	.63
.90	1.00	1.00	1.00	.99	.97	.94	.88	.82	.74	.67	.63
.95	1.00	1.00	1.00	1.00	.99	.97	.92	.84	.75	.67	.63
1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.86	.75	.67	.63

Proportion of Employees Considered Satisfactory = .70
Selection Ratio

r	.05	.10	.20	.30	.40	.50	.60	.70	.80	.90	.95
.00	.70	.70	.70	.70	.70	.70	.70	.70	.70	.70	.70
.05	.73	.73	.72	.72	.72	.71	.71	.71	.71	.70	.70
.10	.77	.76	.75	.74	.73	.73	.72	.72	.71	.71	.70
.15	.80	.79	.77	.76	.75	.74	.73	.73	.72	.71	.71
.20	.83	.81	.79	.78	.77	.76	.75	.74	.73	.71	.71
.25	.86	.84	.81	.80	.78	.77	.76	.75	.73	.72	.71
.30	.88	.86	.84	.82	.80	.78	.77	.75	.74	.72	.71
.35	.91	.89	.86	.83	.82	.80	.78	.76	.75	.73	.71
.40	.93	.91	.88	.85	.83	.81	.79	.77	.75	.73	.72
.45	.94	.93	.90	.87	.85	.83	.81	.78	.76	.73	.72
.50	.96	.94	.91	.89	.87	.84	.82	.80	.77	.74	.72
.55	.97	.96	.93	.91	.88	.86	.83	.81	.78	.74	.72
.60	.98	.97	.95	.92	.90	.87	.85	.82	.79	.75	.73
.65	.99	.98	.96	.94	.92	.89	.86	.83	.80	.75	.73
.70	1.00	.99	.97	.96	.93	.91	.88	.84	.80	.76	.73
.75	1.00	1.00	.98	.97	.95	.92	.89	.86	.81	.76	.73
.80	1.00	1.00	.99	.98	.97	.94	.91	.87	.82	.77	.73
.85	1.00	1.00	1.00	.99	.98	.96	.93	.89	.84	.77	.74
.90	1.00	1.00	1.00	1.00	.99	.98	.95	.91	.85	.78	.74
.95	1.00	1.00	1.00	1.00	1.00	.99	.98	.94	.86	.78	.74
1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.88	.78	.74

Proportion of Employees Considered Satisfactory = .80
Selection Ratio

r	.05	.10	.20	.30	.40	.50	.60	.70	.80	.90	.95
.00	.80	.80	.80	.80	.80	.80	.80	.80	.80	.80	.80
.05	.83	.82	.82	.82	.81	.81	.81	.81	.81	.80	.80
.10	.85	.85	.84	.83	.83	.82	.82	.81	.81	.81	.80
.15	.88	.87	.86	.85	.84	.83	.83	.82	.82	.81	.81
.20	.90	.89	.87	.86	.85	.84	.84	.83	.82	.81	.81
.25	.92	.91	.89	.88	.87	.86	.85	.84	.83	.82	.81
.30	.94	.92	.90	.89	.88	.87	.86	.84	.83	.82	.81
.35	.95	.94	.92	.90	.89	.89	.87	.85	.84	.82	.81
.40	.96	.95	.93	.92	.90	.89	.88	.86	.85	.83	.82
.45	.97	.96	.95	.93	.92	.90	.89	.87	.85	.83	.82
.50	.98	.97	.96	.94	.93	.91	.90	.88	.86	.84	.82
.55	.99	.98	.97	.95	.94	.92	.91	.89	.87	.84	.82
.60	.99	.99	.98	.96	.95	.94	.92	.90	.87	.84	.83
.65	1.00	.99	.98	.97	.96	.95	.93	.91	.88	.85	.83
.70	1.00	1.00	.99	.98	.97	.96	.94	.92	.89	.85	.83
.75	1.00	1.00	1.00	.99	.98	.97	.95	.93	.90	.86	.83
.80	1.00	1.00	1.00	1.00	.99	.98	.96	.94	.91	.87	.84
.85	1.00	1.00	1.00	1.00	1.00	.99	.98	.96	.92	.87	.84
.90	1.00	1.00	1.00	1.00	1.00	1.00	.99	.97	.94	.88	.84
.95	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.99	.96	.89	.84
1.00	1 00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.89	.84

Proportion of Employees Considered Satisfactory = .90
Selection Ratio

r	.05	.10	.20	.30	.40	.50	.60	.70	.80	.90	.95
.00	.90	.90	.90	.90	.90	.90	.90	.90	.90	.90	.90
.05	.92	.91	.91	.91	.91	.91	.91	.90	.90	.90	.90
.10	.93	.93	.92	.92	.92	.91	.91	.91	.91	.90	.90
.15	.95	.94	.93	.93	.92	.92	.92	.91	.91	91	.90
20	.96	.95	.94	.94	.93	.93	.92	.92	.91	91	.90
25	.97	.96	.95	.95	.94	.93	.93	.92	.92	.91	.91
.30	.98	.97	.96	.95	.95	.94	.94	.93	.92	.91	.91
.35	.98	.98	.97	.96	.95	.95	.94	.93	.93	.92	.91
.40	.99	.98	.98	.97	.96	.95	.95	.94	.93	.92	91
.45	.99	.99	.98	.98	.97	.96	.95	.94	.93	92	.91
.50	1.00	.99	.99	.98	.97	.97	.96	.95	.94	.92	.92
.55	1.00	1.00	.99	.99	.98	.97	.97	.96	.94	.93	.92
.60	1.00	1.00	.99	.99	.99	.98	.97	.96	.95	.93	.92
.65	1.00	1.00	1.00	.99	.99	.98	.98	.97	.96	.94	.92
.70	1.00	1.00	1.00	1.00	.99	.99	.98	.97	.96	.94	.93
.75	1.00	1.00	1.00	1.00	1.00	.99	.99	.98	97	.95	.93
.80	1.00	1.00	1.00	1.00	1.00	1.00	.99	.99	.97	.95	.93
.85	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.99	.98	.96	.94
.90	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.99	.97	.94
.95	1.00	1.00	1.00	1 00	1.00	1.00	1.00	1.00	1 00	.98	.94
1 00	1.00	1.00	1.00	1 00	1.00	1.00	1.00	1.00	1.00	1.00	.95

Appendix C

Publishers of Tests[1]

1. Intelligence Tests

Title	*Publisher*[2]
Adaptability Test	Science Research Associates, 228 S. Wabash Ave. Chicago, Illinois
Army Alpha, Bregman's Revision	Psychological Corporation, 522 Fifth Avenue, New York City
Army Alpha, Forms 5 and 7	Psychological Corporation
Army Alpha, Form 6	Psychological Corporation
Army Alpha, Nebraska Revision	Sheridan Supply Company, P. O. Box 837, Beverly Hills, Cal.
Army Alpha, Schrammel-Brannan	Bureau of Education Measurements, Kansas State Teachers College, Emporia, Kansas
Benge Test of General Knowledge	Management Service Company, 3136 N. 24th Street, Philadelphia, Pa.
Bureau Test VI	Psychological Corporation
Detroit General Aptitudes Examination	Public School Publishing Company, Bloomington, Ill.
Henmon-Nelson Tests of Mental Ability	Psychological Corporation
O'Rourke General Classification Test	The Psychological Institute, 3506 Patterson St., N. W., Washington, D. C.
Otis S-A Test of Mental Ability	World Book Company, Yonkers, N. Y.
Pressey Senior Classification Test	Public School Publishing Company
Pressey Senior Verifying Test	Public School Publishing Company
Purdue Industrial Training Classification Test, Forms A and B	Science Research Associates

[1] This summary is reproduced in part from *Experience with Employment Tests*, Studies in Personnel Policy, No. 32, National Industrial Conference Board, Inc., 247 Park Avenue, New York, N. Y.

[2] The address of the publisher is given in full only the first time the company appears in this list.

Title	*Publisher*[1]
Roback Mentality Tests	C. H. Stoelting, 424 N. Homan Avenue, Chicago, Ill.
Scott Company Mental Alertness Test	C. H. Stoelting
Stanford Scientific Aptitude Test, D. L. Zyve	Stanford University Press, Stanford University, Cal.
Terman Group Test of Mental Ability	Psychological Corporation
Vocational Aptitude Examination (Cleeton & Mason)	Psychological Corporation
Wonderlic Personnel Test	E. F. Wonderlic, 919 N. Michigan Ave., Chicago, Ill.

2. Clerical Tests

Benge Clerical Test D	Management Service Company
Benge's Stenogauge	Psychological Corporation
Bennett's Stenographic Test	Psychological Corporation
Blackstone Typewriting Test	Psychological Corporation
Clapp-Young Arithmetic Test	Psychological Corporation
Clem Senior Typewriting Test	Public School Publishing Company
Cole Guidance Examinations in the Fundamental Vocabulary	Public School Publishing Company
Detroit Clerical Aptitudes Examination	Public School Publishing Company
E.R.C. Stenographic Aptitude Test	Science Research Associates
Kimberly Clark Typing Ability Analysis	Science Research Associates
Link's Context Test	C. H. Stoelting
Markham English Vocabulary Tests	Public School Publishing Company
Michigan Vocabulary Profile Test Form BM	World Book Company
N.I.I.P. Clerical Test	Psychological Corporation
Minnesota Test for Clerical Workers	Psychological Corporation
O'Connor's English Vocabulary, Work-sample 95, Form EA	Stevens Institute of Technology, Hoboken, New Jersey
O'Rourke Clerical Aptitude Test, Junior Grade	Psychological Institute
Otis Arithmetic Test	World Book Company
Pressey Diagnostic Test in Eng. Comp.	Public School Publishing Company
Rogers' Stenographic and Typing Tests	C. H. Stoelting
Schorling—Clark Potter Arithmetic Test	Psychological Corporation
Scott Company File Clerk Test	C. H. Stoelting
Shellow's Intelligence Test for Stenographers and Typists	C. H. Stoelting
SRH Tests of Clerical Aptitude	Science Research Associates
SRA Test of Language Skill	Science Research Associates
SRA Test of Typing Skill	Science Research Associates
SRA Test of Dictation Skill	Science Research Associates

[1] The address of the publisher is given in full only the first time the company appears in the list.

Title	*Publisher*[1]
Thurstone Examination in Clerical Work, Form A	World Book Company
Thurstone Examination in Typing, Form A	Psychological Corporation
Tressler Minimum Essentials Test (English)	Public School Publishing Company

3. Dexterity and Manipulative Tests

MacQuarrie Test for Mechanical Ability	California Test Bureau 3836 Beverly Boulevard Los Angeles, California
Minnesota Rate of Manipulation Test	Psychological Corporation
O'Connor Finger Dexterity Test	Human Engineering Laboratory Stevens Institute of Technology Hoboken, New Jersey
O'Connor Tweezer Dexterity Test	Human Engineering Laboratory
Purdue Hand Precision Test	Lafayette Instrument Company Lafayette, Indiana
Purdue Pegboard	Science Research Associates
Stanford Motor Skills Unit	C. H. Stoelting

4. Mechanical Tests

Bennett's Test of Mechanical Comprehension, Forms A and AA	Psychological Corporation
Detroit Manual Ability	C. H. Stoelting
Detroit Mechanical Aptitudes Examination, Form A	Public School Publishing Company
Industrial Training Classification Test	Science Research Associates
Kent-Shakow Industrial Form Boards	C. H. Stoelting
Mellenbruch Tests of Mechanical Aptitude	Science Research Associates
Minnesota Mechanical Assembly Test	Psychological Corporation
Minnesota Paper Form Board, Series B, Revised Series AA	Science Research Associates
Minnesota Spatial Relations Test	Psychological Corporation
O'Connor Wiggly Block	Stevens Institute
O'Rourke Mechanical Aptitude Test, Junior Grade	Psychological Corporation
Purdue Mechanical Adaptability Test	Division of Applied Psychology Purdue University Lafayette, Indiana
Stenquist Mechanical Aptitude Test	Psychological Corporation
SRA Tests of Mechanical Aptitude	Science Research Associates

5. Personality Tests

Allport Scale of Values Test	Houghton Mifflin Company, 2 Park Street, Boston, Mass.

[1] The address of the publisher is given in full only the first time the company appears in the list.

Title	*Publisher*[1]
A-S Reaction (Beckman's Revision)	Psychological Corporation
Bell's Adjustment Inventory	Stanford University Press
Benge Interest Test	Management Service Co.
Bernreuter Personality Inventory	Psychological Corporation
California Test of Personality, Secondary Form A	California Test Bureau
Guilford Series of Personality Tests	Sheridan Supply Company Beverly Hills, California
Humm-Wadsworth Temperament Scale	D. G. Humm Personnel Service, Los Angeles, California
Inventory of Factors S T D C R	Sheridan Supply Company
Keeler Polygraph	Associated Research Inc., 431 South Dearborn St., Chicago, Illinois
Laird's I-E Test	Psychological Corporation
Minnesota Multiphasic Personality Inventory	University of Minnesota Press Minneapolis, Minnesota
Psycho-Somatic Inventory	Psychological Corporation
Root's New I-E Test	Psychological Corporation
Thurstone Personality Schedule	University of Chicago

6. Interest Tests

Brainard-Stewart Specific Interest Inventories	Psychological Corporation
Cleeton Vocational Interest Inventory	Psychological Corporation
Kuder Preference Record	Science Research Associates
Le Suer Occupational Interest Blank	Psychological Corporation
Link's Personality Quotient Test	Psychological Corporation
Manson Occupational Interest Blank for Women	Psychological Corporation
Minnesota Interest Analysis	Psychological Corporation
Primary Business Interests Test	Science Research Associates
Strong Vocational Interest Blank for Men	Stanford University Press
Thurstone Vocational Interest Schedule	Psychological Corporation

7. Trade Tests

Can You Read a Working Drawing?	Science Research Associates
Can You Read a Micrometer?	Science Research Associates
Can You Read a Scale?	Science Research Associates
Purdue Vocational Tests	
1. Machine Shop and Machine Operation	Science Research Associates
(Total score for general machinist and sub-scores for operators of lathe, shaper and planer, grinder, milling machine, and bench workers)	

[1] The address of the publisher is given in full only the first time the company appears in the list.

Title	*Publisher*[1]
2. Electricity	Science Research Associates
3. Blueprint Reading	Science Research Associates
Purdue Industrial Mathematics Test	Division of Applied Psychology Purdue University Lafayette, Indiana

8. Industrial Vision Tests

Ortho-Rater Visual Classification Tests	Bausch and Lomb Optical Company Rochester, New York
Site-Screener	American Optical Company Cambridge, Massachusetts
Keystone Visual Safety Tests	Keystone View Company Meadville, Pennsylvania

[1] The address of the publisher is given in full only the first time the company appears in the list.

INDEX

Index

B

C

G

H

Q

R

S

U